UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE BIOLOGIA VEGETAL

Clinical and molecular characterization of Portuguese patients with a clinical diagnosis of MODY

João Paulo de Medeiros Gomes de Mafra

Mestrado em Biologia Molecular e Genética

Dissertação orientada por:
Mafalda Bourbon

2017

Agradecimentos

À Mafalda, pela oportunidade de ganhar experiência em ambiente laboratorial e de participar num congresso internacional. À Ana e Catarina pelo apoio inicial a lidar com assuntos que me eram estranhos. À Chora pela ajuda com os *in silico* e à Cibelle pelo sorriso e boa disposição. À Gisela pela colaboração no MLPA.

À Sílvia pelo bom humor e leveza. À Joana Costa pelas picardias e conversas futebolísticas. À Andreia por ser um doce de pessoa. Ao Niccolò pelo escape e pelo *man time*. À Inês pelo companheirismo. À Russo e Sara pela ajuda na integração à FCUL.

À Canilho por uma amizade que ultrapassa o laboratório. À Rita por me dar a mão quando caí. Ao Pedro pela irmandade e sintonia intelectual. Aos que ficam sem nome mas que não me esquecem, que acreditam em mim, nem eu sei bem porquê, mas que também não esqueço e muito estimo. Aos irmãos que por vezes sinto que perdi.

Aos meus pais, que me apoiam mais do que mereço.

Obrigado.

Resumo

A diabetes mellitus, ou simplesmente diabetes, pode ser definida como um conjunto complexo de perturbações crónicas de cariz metabólico caracterizadas por hiperglicémia. A diabetes tipo MODY, do inglês *Maturity-onset diabetes of the young*, é uma forma monogénica de diabetes. Inicialmente descrita em 1974 por Tattersall, a diabetes tipo MODY engloba um grupo de fenótipos heterogéneos, clinica e geneticamente, caracterizados por alterações no funcionamento normal das células beta do pâncreas e por um padrão de hereditariedade autossómico dominante. De uma forma geral, a diabetes tipo MODY tem um perfil não-insulino dependente e manifesta-se em crianças e indivíduos jovens, sendo tipicamente diagnosticada antes dos 25 anos.

A diabetes tipo MODY aparenta ser rara, estimando-se que seja responsável por 0,6-2% dos casos de diabetes na Europa. No entanto, é frequente esta forma de diabetes ser equivocamente diagnosticada como diabetes tipo 1 ou tipo 2, pelo que a sua prevalência real deverá ser superior. Um dos grandes trunfos no combate a este subdiagnóstico da diabetes tipo MODY são os testes genéticos. Desde a década de 1990, foram 13 os genes associados à MODY. Mutações em heterozigotia nos genes *GCK* e *HNF1A* são as causas mais frequentes de diabetes tipo MODY, correspondendo a cerca de 70% dos casos. Logo a seguir estão as mutações em heterozigotia nos genes *HNF4A* (*hepatocyte nuclear factor 4 alpha*) e *HNF1B*, que correspondem a cerca de 15% dos casos.

O gene *GCK*, localizado no cromossoma 7 (7p13), codifica o enzima glicocinase (*glucokinase* - GCK), também conhecido como hexocinase IV. Este enzima monomérico possui três isoformas - a isoforma 1, presente nas células beta, e as 2 e 3, presentes no fígado - e, no interior das células, atua como um sensor do nível de glicose. Nos hepatócitos, este enzima intervém no desencadear da glicólise e glicogénese, enquanto facilita a exocitose de insulina nas células beta. A diabetes tipo MODY, subtipo GCK, resulta de mutações de perda de função em heterozigotia no gene *GCK*, diminuindo a atividade do enzima, e caracteriza-se por uma hiperglicémia moderada, assintomática e não progressiva que se manifesta desde o nascimento. Estas mutações acabam por diminuir a quantidade de glicogénio sintetizado e impedem a normal libertação de insulina, ao aumentar a concentração mínima de glicose necessária à sua secreção.

O gene *HNF1A*, localizado no cromossoma 12 (12q24.31), codifica um fator de transcrição (*hepatocyte nuclear factor 1 alpha* - HNF1A) homodimérico. Este fator de transcrição possui três isoformas, A, B e C. As três estão presentes no fígado, rins, pâncreas e intestinos mas a primeira predomina no fígado, rins e pâncreas fetal, enquanto a isoforma B predomina no pâncreas adulto. O HNF1A integra uma complexa rede de fatores de transcrição, desempenhando um papel regulador na expressão de diversos genes durante o desenvolvimento embrionário. No fígado, o HNF1A regula a expressão de vários genes hepáticos, como o gene que codifica para a albumina. Nas células beta, intervém na expressão de insulina e na proliferação e morte celular. A diabetes tipo MODY, subtipo HNF1A, resulta de mutações em heterozigotia no gene *HNF1A*. Estas mutações podem ter diversos efeitos, reduzindo a secreção de insulina em resposta a glicose e aminoácidos, e alterando a expressão de genes envolvidos no transporte (*GLUT2*) e metabolismo de glicose, afetando processos como a glicólise, gluconeogénese e derivação de aminoácidos para o ciclo de Krebs. O subtipo HNF1A está associado a defeitos na proliferação das células beta e caracteriza-se por uma incapacidade progressiva na secreção de insulina, que não acompanha o aumento de glicose em circulação, dando origem a uma hiperglicémia mais grave que o subtipo GCK.

O gene *HNF1B*, localizado no cromossoma 17 (17q12), codifica um fator de transcrição (*hepatocyte nuclear factor 1beta* - HNF1B) que atua como homodímero ou heterodímero, com HNF1A. O gene

HNF1B produz três isoformas - 1, 2 e 3 - e é expresso no timo, pulmão, rim, fígado, pâncreas, estômago, intestino e trato genital. Este fator de transcrição, que integra a mesma rede regulatória que HNF1A, actua no desenvolvimento embrionário do pâncreas e do rim. A diabetes tipo MODY, subtipo HNF1B, resulta de mutações em heterozigotia no gene *HNF1B* e em cerca de 50% dos casos caracteriza-se por uma combinação de resistência à insulina e disfunção das células beta, exibindo, à semelhança do subtipo HNF1A, uma incapacidade na secreção de insulina em resposta a concentrações crescentes de glicose. Mutações neste gene podem também resultar em anomalias extra pancreáticas e afetar tecidos como o trato genital ou o fígado, sendo quistos renais o fenómeno mais observado.

Neste estudo, foram recolhidas informações clínicas acerca de 39 indivíduos, 24 probandos e 15 familiares, através de questionários enviados por vários médicos de diferentes hospitais portugueses. Com base nestas informações, como glicémia em jejum, prova de tolerância à glicose oral e HbA1c, e no diagnóstico clínico de diabetes tipo MODY, foram-nos referenciados indivíduos com uma história familiar de diabetes que evidencie um padrão hereditário dominante; sem autoanticorpos pancreáticos; com início de sintomas antes dos 25 anos; e índice de massa corporal maioritariamente normal. Estes 24 probandos foram estudados por sequenciação de Sanger para os genes *GCK* e *HNF1A*, tendo sido também efetuada a pesquisa de inserções e deleções através da técnica MLPA (*Multiplex Ligation-dependent Probe Amplification*).

Entre substituições pontuais e pequenas deleções, a sequenciação de Sanger detetou 46 variantes genéticas, 19 no gene *GCK* e 27 no gene *HNF1A*. Entre estas, seis podem ser classificadas como patogénicas ou provavelmente patogénicas, sendo quatro no gene *GCK* e duas no gene *HNF1A*. As variantes patogénicas ou provavelmente patogénicas detetadas no gene *GCK* foram c.364C>T (p.(Leu122Phe)), uma alteração missense no exão 4 que cosegrega com diabetes e está associada à diabetes tipo MODY; c.579+1_579+33del, uma deleção de 33 pares de base no intrão 5 que elimina um local de *splicing* e também está associada à diabetes tipo MODY; c.766G>A (p.Glu256Lys), uma alteração missense no exão 7 que induz alterações conformacionais no enzima GCK, reduzindo a sua capacidade de ligação à glicose e subsequente atividade catalítica; e, finalmente, c.1268T>A (p.(Phe423Tyr)), uma alteração missense no exão 10 que cosegrega com diabetes e está associada à diabetes tipo MODY. Estas variantes foram encontradas num total de quatro probandos e cinco familiares. No gene *HNF1A*, as variantes patogénicas detetadas foram c.814C>T (p.(Arg272Cys)), uma alteração missense no exão 4 que cosegrega com diabetes e impede o fator de transcrição HNF1A de se ligar ao DNA, perdendo assim a sua atividade reguladora na expressão genética; e c.872del (p.(Pro291Glnfs*51)), uma deleção também localizada no exão 4 que provoca uma alteração na grelha de leitura, cosegrega com diabetes e possivelmente resulta de um fenómeno de *splippage* aquando da replicação de DNA, prevendo-se que resulte numa proteína truncada. Estas variantes foram encontradas num total de três probandos e três familiares.

A técnica MLPA foi aplicada na investigação de inserções e deleções em 17 probandos. Três destes probandos geraram resultados sem qualidade, sendo necessárias novas amostras para eventual repetição. Noutros 11 nenhuma inserção ou deleção foi detetada. Nos restantes três probandos foram detetadas duas deleções patogénicas: deleção em heterozigotia dos exões 5 a 8 do gene *GCK* (c.484-?_1019+?del) num probando; e deleção em heterozigotia do gene *HNF1B* (c.1-?_1674+?del) em dois probandos. A deleção no gene *GCK* deverá produzir uma proteína não funcional para gerar o fenótipo MODY, subtipo GCK. No entanto, não foi encontrada qualquer informação sobre esta deleção ou o seu efeito na proteína, pelo que poderemos estar na presença de uma nova mutação. Quanto à deleção do gene *HNF1B*, estas são frequentes e existem várias fontes que apontam deleções completas do gene como patogénicas, que muitas vezes incluem outros genes na mesma região (17q12). Um dos

probandos aparenta não ter história familiar de diabetes e, sendo mutações *de novo* frequentes, é possível que estejamos na presença de uma. No entanto, não haviam amostras de familiares para fazer estudos de cosegregação nestes dois probandos. A hemizigotia provocada por esta deleção deve resultar em MODY, subtipo HNF1B, por haploinsuficiência. Nenhum destes dois probandos aparenta ter anomalias renais.

No total, entre 24 probandos, este estudo identificou 10 indivíduos com MODY, cinco do subtipo GCK, três do subtipo HNF1A e dois do subtipo HNF1B, realçando assim a importância de um diagnóstico correto, com recurso a ferramentas de genética molecular, uma vez que estes indivíduos tinham diagnósticos de diabetes tipo 1 ou tipo 2. Dos restantes 14 probandos sem qualquer variante patogénica/provavelmente patogénica, foi detetada pelo menos uma variante associada a diabetes tipo 2. A utilização destas técnicas no âmbito do diagnóstico genético permite estabelecer o diagnóstico correcto, com implicações para a terapêutica a administrar e qualidade de vida dos utentes dos serviços de saúde.

Palavras-chave: Diabetes; MODY; GCK; HNF1A; HNF1B

Abstract

Initially described by Tattersall in 1974, maturity-onset diabetes of the young (MODY) is a form of early onset monogenic diabetes characterized by clinically heterogeneous phenotypes with autosomal dominant inheritance that generally result in β-cell dysfunction. Often misdiagnosed, MODY accounts for 0.6-2% of diabetes cases in Europe and 13 genes have been implicated in this form of diabetes, with heterozygous mutations in *GCK* and *HNF1A* being the most common etiologies, followed by heterozygous mutations in *HNF4A* and *HNF1B.*

Glucokinase (GCK) is a monomeric enzyme that acts as a cellular glucose sensor, playing a role in glycogenesis and glycolysis in hepatocytes, and in insulin release in β-cells. Heterozygous loss-of-function mutations in *GCK* decrease enzymatic activity and result in GCK-MODY, characterized by asymptomatic mild stable hyperglycemia present from birth. HNF1A-MODY is associated with impaired β-cell proliferation and is characterized by a progressive insulin secretory defect, resulting in a more severe hyperglycemia. This form of MODY is caused by heterozygous mutations in *HNF1A*, a gene that codes hepatocyte nuclear factor 1 alpha, a homodimeric transcription factor that regulates gene expression in embryonic development and plays a role in insulin expression and β-cell proliferation and death. Hepatocyte nuclear factor 1 beta (HNF1B) is also a transcription factor that acts as a homodimer, or a heterodimer with HNF1A. HNF1B is part of the same regulatory network as HNF1A and regulates embryonic pancreatic development, also playing a role in nephron development. Heterozygous *HNF1B* mutations result in HNF1B-MODY, characterized by a combination of β-cell dysfunction and insulin resistance in approximately 50% of mutation carriers. Like in HNF1A-MODY, insulin secretion is compromised as glucose concentrations rise. Renal cysts are common.

Clinical data on 24 probands and 15 relatives was collected through questionnaires sent by physicians. Subjects with a clinical diagnosis of MODY - family history of diabetes consistent with dominant inheritance pattern, no pancreatic autoantibodies, onset before 25 years of age and typically lean - were referred to this study by physicians for genetic screening via Sanger sequencing (for *GCK* and *HNF1A*) and Multiplex Ligation-dependent Probe Amplification (MLPA).

Between point substitutions and small deletions, Sanger sequencing detected 46 variants, 19 in *GCK* and 27 in *HNF1A*. Six of these were pathogenic or likely pathogenic, four in *GCK* and two in *HNF1A*. *GCK* pathogenic or likely pathogenic variants were c.364C>T (p.(Leu122Phe)), in exon 4; c.579+1_579+33del, in intron 5, abolishing the donor splice site; c.766G>A (p.Glu256Lys), in exon 7, which induces conformational changes, decreasing glucose binding and catalytic activity; and c.1268T>A (p.(Phe423Tyr)), in exon 10. These variants were detected in four probands and five relatives. *HNF1A* pathogenic variants were c.814C>T (p.(Arg272Cys)), which renders HNF1A unable to bind DNA and exert its transactivating activity; and c.872del (p.(Pro291Glnfs*51)), predicted to produce a truncated protein and likely the result of replication slippage. Both variants are located in exon 4. These variants were detected in three probands and three relatives.

MLPA analysis detected two pathogenic deletions: a heterozygous *GCK* exons 5 through 8 deletion (c.484-?_1019+?del) in one proband; and a heterozygous *HNF1B* deletion (c.1-?_1674+?del) in two probands. The seemingly unreported heterozygous *GCK* exons 5 through 8 deletion, which should result in a null variant that decreases overall enzymatic activity, could be a novel mutation. Heterozygous *HNF1B* deletions should result in HNF1B-MODY via haploinsufficiency. One proband had no apparent family history but *de novo* mutations are frequent. Both probands had no apparent renal abnormalities.

Among 24 probands, this study identified a total of 10 MODY cases, with five being GCK-MODY, three being HNF1A-MODY and two being HNF1B-MODY. Of the remaining 14 probands without any detected pathogenic/likely pathogenic variants, 11 had at least one type 2 diabetes associated variant. Hence, this study highlights the importance of genetic diagnosis in patients with diabetic phenotypes consistent with MODY, as a correct diagnosis impacts patient treatment and quality of life.

Key words: Diabetes; MODY; GCK; HNF1A; HNF1B

Table of contents

Agradecimentos I
Resumo II
Abstract V
Abbreviations, acronyms and symbols X
1. Introduction 1
1.1 MODY 1
1.2 Differentiating MODY from T1DM and T2DM 2
1.3 *GCK* and GCK-MODY 3
1.4 *HNF1A* and HNF1A-MODY 5
1.5 *HNF4A* and HNF4A-MODY 8
1.6 *HNF1B* and HNF1B-MODY 10
1.7 Rare types of MODY 11
2. Materials and Methods 12
2.1 Inclusion criteria 12
2.2 DNA extraction 12
2.3 DNA quantification and quality screen 12
2.4 Primer design, checks and preparation 12
2.5 PCR optimization 13
2.6 PCR reactions 13
2.7 Agarose gel electrophoresis 13
2.8 PCR product purification and Sanger sequencing reaction 13
2.9 Sequence analysis and variant interpretation 14
2.10 MLPA 14
3. Results 15
3.1 Subjects under study 15
3.1.1 Number of subjects, sex, ethnicity, age and age at diagnosis 15
3.1.2 Family history of diabetes and renal disease 15
3.1.3 Vascular complications 15
3.1.4 Birth weight and BMI 15
3.1.5 FPG 16
3.1.6 OGTT 16
3.1.7 HbA1c 16
3.1.8 Pancreatic autoantibodies 16
3.1.9 Treatment type (diet, OHA and insulin) 16

3.2 Molecular analysis: Sanger sequencing and MLPA 17
3.2.1 Sanger sequencing analysis: *GCK* 17
3.2.2 Sanger sequencing analysis: *HNF1A* 20
3.2.3 MLPA analysis 24
4. Discussion 26
5. References 31
6. Supplementary material 39
6.1 Comparative analysis of MODY, T1DM and T2DM 39
6.2 Online tools 40
6.3 Interpreting *in silico* mutation predictions 40
6.4 Primer sequences, PCR master mix and cycling program 40
6.4.1 Primer sequences 40
6.4.2 PCR master mix and cycling program 41
6.5 Sanger sequencing mix and cycling program 42
6.6 Sequence analysis and pregap module configuration 42
6.7 Detected sequence variants 43

Tables and figures

Table 6.1 Differentiating MODY from T1DM and T2DM 39
Table 6.2: Primer data 40
Table 6.3: Standard PCR mix and cycling program 41
Table 6.4: Sequencing mix and cycling program 42
Table 6.5: Detected sequence variants 43
Figure 1.1 *HNF1A* structure 7
Figure 1.2 *HNF4A* gene and isoform structures 9
Figure 1.3 *HNF1B* structure 11
Figure 3.1 Detected sequence variants 17
Figure 3.2 *GCK* sequence variants location 17
Figure 3.3 *GCK* sequence variants and affected subjects 18
Figure 3.4 *GCK* c.364C>T proband chromatogram and pedigree 18
Figure 3.5 *GCK* c.579+1_579+33del proband chromatogram and pedigree 19
Figure 3.6 *GCK* c.766G>A proband chromatogram and pedigree 19
Figure 3.7 *GCK* c.1268T>A proband chromatogram and pedigree 20
Figure 3.8 *HNF1A* sequence variants location 20
Figure 3.9 *HNF1A* sequence variants and affected subjects 21
Figure 3.10 *HNF1A* c.814C>T proband (14-014) chromatogram and pedigree 21
Figure 3.11 *HNF1A* c.814C>T proband (14-015) chromatogram and pedigree 22
Figure 3.12 *HNF1A* c.872del proband chromatogram and pedigree 22
Figure 3.13 *HNF1A* c.1720G>A proband chromatogram and pedigree 23
Figure 3.14 *HNF1A* c.79A>C proband chromatograms 23
Figure 3.15 *HNF1A* c.293C>T and c.1460G>A proband chromatograms 24
Figure 3.16: MLPA results for proband 15-014 25
Figure 3.17: MLPA results for probands 13-006 (A) and 14-013 (B) 25

Abbreviations, acronyms and symbols

°C	Celsius Degree
µL	Microlitre (10^{-6} l)
µM	Micromolar (10^{-6} M)
ABCC8	ATP Binding Cassette subfamily C member 8
ACMG	American College of Medical Genetics and Genomics
ADP	Adenosine Diphosphate
AE	At Enrollment
AFR	African population
Ala	Alanine
ALT	Alanine Aminotransferase
AMP	Association for Molecular Pathology
apo A1	Apolipoprotein A1
apo A2	Apolipoprotein A2
apo B	Apolipoprotein B
apo CIII	Apolipoprotein CIII
Arg	Arginine
Asn	Asparagine
Asp	Aspartic Acid
ATP	Adenosine Triphosphate
BE	Before Enrollment
BLAT	BLAST-like Alignment Tool
BLK	B Lymphocyte Kinase, gene
BMI	Body Mass Index
bp	Base Pair
Ca^{2+}	Calcium ion
CEL	Carboxyl Ester Lipase, gene
CNV	Copy Number Variation
Cys	Cysteine
del	Deletion
DKA	Diabetic Ketoacidosis
DNA	Deoxyribonucleic Acid
dNTP	Deoxynucleotide Triphosphates
EDTA	Ethylenediamine Tetraacetic Acid
EUR	European population
EVS	Exome Variant Server
ExAC	Exome Aggregation Consortium
F	Forward (primer)
FPG	Fasting Plasma Glucose
fs	Frameshift
g	Gram
G6P	Glucose-6-phosphate
GADA	Glutamic Acid Decarboxylase Antibodies
GCK	Glucokinase, gene
GDM	Gestational Diabetes
GGT	Gamma-glutamyl transferase
GKRP	GCK Regulatory Protein
Gln	Glutamine
Glu	Glutamic Acid
GLUT2	Glucose Transporter 2
Gly	Glycine
HbA1c	Glycated Hemoglobin
HDL	High Density Lipoprotein
HGMD	Human Gene Mutation Database

HNF	Hepatocyte Nuclear Factor
HNF1A	Hepatocyte Nuclear Factor 1 Alpha, gene
HNF1B	Hepatocyte Nuclear Factor 1 Beta, gene
HNF4A	Hepatocyte Nuclear Factor 4 Alpha, gene
i.e.	*id est*
ICA	Islet Cell Antibodies
IGT	Impaired Glucose Tolerance
Ile	Isoleucine
INS	Insulin, gene
kbp	Kilo (10^3) Base Pair
KCNJ11	Potassium Voltage-gated Channel subfamily J member 11, gene
kg	Kilogram
kg/m^2	Kilogram per Square Metre
KLF11	Kruppel Like Factor 11, gene
K_m	Michaelis constant
L	Litre
LDL	Low Density Liproprotein
Leu	Leucine
Lys	Lysine
MAF	Minor Allele Frequency
mg/dL	Milligram (10^{-3} g) per Decilitre (10^{-1} L)
$MgCl_2$	Magnesium Chloride
min	Minute
mL	Millilitre (10^{-3} L)
MLPA	Multiplex Ligation-dependent Probe Amplification
mM	Millimolar (10^{-3} M)
mmol/L	Millimol (10^{-3} mol) per Litre
MODY	Maturity-onset Diabetes of the Young
mRNA	Messenger Ribonucleic Acid
NA	Not Applicable
NCBI	National Center for Biotechnology Information
ND	No Data available
NEUROD1	Neuronal Differentiation 1, gene
NH4	Ammonium
NMD	Nonsense-mediated Decay
OGTT	Oral Glucose Tolerance Test
OHA	Oral Hypoglycemic Agent
P1	Hepatic Promoter (*HNF4A*)
P2	Pancreatic Promoter (*HNF4A*)
PAX4	Paired Box 4, gene
PCR	Polymerase Chain Reaction
PDX1	Pancreatic and Duodenal Homeobox 1, gene
Phe	Phenylalanine
PNDM	Permanent Neonatal Diabetes Mellitus
PolyPhen-2	Polymorphism Phenotyping v2
Pro	Proline
PROVEAN	Protein Variation Effect Analyzer
R	Reverse (primer)
RCAD	Renal Cysts and Diabetes
RefSeq	Reference Sequence
RNA	Ribonucleic Acid
sec	Second
Ser	Serine
SGLT-2	Sodium Glucose Cotransporter 2
SIFT	Sorting Intolerant From Tolerant

SNP	Single Nucleotide Polymorphism
T1DM	Type 1 Diabetes Mellitus
T2DM	Type 2 Diabetes Mellitus
TBE	Tris-Borate-EDTA
TCA	Tricarboxylic Acid
Thr	Threonine
Tyr	Tyrosine
UCSC	University of California Santa Cruz
UK	United Kingdom
UTR	Untranslated Region
UV	Ultraviolet
Val	Valine
V_{max}	Maximal Velocity
WD	Without Diagnosis
ρmol/L	Picomol (10^{-12} mol) per litre

1. Introduction

1.1 MODY

Maturity-onset diabetes of the young (MODY) is a form of early onset monogenic diabetes (1). It was first described in 1974 by Tattersall, who reported on three families with seemingly dominantly inherited mild non-progressive diabetes diagnosed in their teen years or early twenties (2). Currently, MODY describes a group of clinically heterogeneous phenotypes of familial diabetes. This group is characterized by monogenic, dominantly inherited autosomal disorders that generally result in β-cell dysfunction. MODY is typically noninsulin-dependent, afflicting children and young adults. As such, its early onset usually leads to a diagnosis before individuals reach 25 years of age (3–5). MODY's heterogeneous clinical profiles result from varying features such as genetic etiology, age of onset, severity of hyperglycemia, extra pancreatic features and treatment modality (4,6).

Though seemingly rare, MODY is the most common form of monogenic diabetes, accounting for an estimated 0.6-2% of all diabetes cases in Europe (4). However, since it is often misdiagnosed as type 1 diabetes (T1DM) or type 2 diabetes (T2DM), actual prevalence should be higher, but no population-based study has been done (3,5). One study found that only 6% of MODY patients were correctly identified clinically, with 36% misdiagnosed as T1DM and 51% as T2DM (7). Another study reported that 25% of subjects with a T2DM diagnosis before 30 years of age actually had MODY, suggesting the use of widened diagnostic testing criteria to bypass this situation (8). This might reflect physicians limited awareness of the condition or cost restrictions in genetic testing. Thus, one cannot understate the importance of a correct diagnosis, which enables physicians to predict the likely clinical course and possibly impacts patient treatment and overall quality of life, as well as the cost effectiveness of the treatment process (3,7). This is attainable through molecular genetic techniques such as direct sequencing, which can diagnose MODY with up to 100% sensitivity by identifying mutations in the gene responsible for the phenotype and defining the associated subtype (3,5,9).

In the 1990s, advancements in molecular genetics allowed identification of some genes responsible for MODY (6). MODY's phenotypes result from heterozygous mutations in genes involved in β-cell development or insulin secretion. Up to 13 genes have been identified and researchers believe many more to be associated, as only 10% to 20% of MODY cases in Japan and China are attributed to known MODY genes (6). Accordingly, investigators are prompted to identify these genes and clarify their role in the pathogenesis of MODY. The 13 genes know to cause MODY are (5,6):

- hepatocyte nuclear factor 4 alpha (*HNF4A*, * 600281), also known as MODY1;
- glucokinase (*GCK*, * 138079), also known as MODY2;
- hepatocyte nuclear factor 1 alpha (*HNF1A*, * 142410), also known as MODY3;
- pancreatic and duodenal homeobox 1 (*PDX1*, * 600733), also known as MODY4;
- hepatocyte nuclear factor 1 beta (*HNF1B*, * 189907), also known as MODY5;
- neuronal differentiation 1 (*NEUROD1*, * 601724), also known as MODY6;
- Kruppel like factor 11 (*KLF11*, * 603301), also known as MODY7;
- carboxyl ester lipase (*CEL*, * 114840), also known as MODY8;
- paired box 4 (*PAX4*, * 167413), also known as MODY9;
- insulin (*INS*, * 176730), also known as MODY10;
- B lymphocyte kinase (*BLK*, * 191305), also known as MODY11;

- ATP binding cassette subfamily C member 8 (*ABCC8*, * 600509), also known as MODY12;
- potassium voltage-gated channel subfamily J member 11 (*KCNJ11*, * 600937), also known as MODY13.

Mutations in *GCK* and *HNF1A* are by far the most common causes of MODY, accounting for nearly 70% of cases (3). The relationship between the two varies according to geographical location, with *GCK* mutations being the most common in Spain, France, Italy (5), Germany, Austria (10), Japan (11) and the Czech Republic (12); and *HNF1A* mutations the most common in Denmark, the UK, the Netherlands (12), China (13) and Korea (14). This variation could be attributed to differences in screening strategies, as countries that seldom perform routine blood glucose tests have higher diagnostic rates for MODY3, as opposed to countries that routinely perform blood glucose tests, which present a higher prevalence of *GCK* mutations (3,5). Extending the focus to include *HNF4A* alongside *GCK* and *HNF1A* mutations, these three account for up to 80% of all cases (3,4). The fourth most common cause of MODY are *HNF1B* mutations, which comprise about 5% of cases (3,5). Each of the remaining nine genes is responsible for up to 1% of cases (5).

Several types of mutation have been identified in *GCK*, *HNF1A* and *HNF4A*, including missense; nonsense; splicing; promoter region variants; frameshifts; small, partial and whole gene deletions; insertions and duplications (3,7,15,16). MODY mutations have a relatively high penetrance (4).

1.2 Differentiating MODY from T1DM and T2DM

Between 50% and 80% of MODY patients are diagnosed with T1DM or T2DM at presentation (5,12). Beyond physician's lack of awareness, due to MODY's rare occurrence, lies a problem posed by overlapping clinical signs between MODY and T1DM/T2DM (5). One study found that 82% of individuals with a molecular diagnosis of MODY first presented with polyuria and polydipsia, and 44% with weight loss complaints, all classical signs of diabetes. The same study also noted that only half the MODY positive subjects had parental history of diabetes, with the same holding true for MODY negative subjects (7).

Typical MODY diagnostic criteria include onset before 25 years of age, family history of diabetes and lack of insulin dependence, and though less than half of individuals meet these criteria, they have displayed high specificity (and low sensitivity) (5,8,17). However, performing genetic tests on any and all individuals, with no regard to specific criteria is not cost-effective and can lead to inappropriate results. So how do we distinguish MODY from T1DM and T2DM? One could argue that MODY and T1DM are easily differentiable based on absence or presence of β-cell autoimmunity, respectively. This is generally true, as highlighted by a UK study (18), but it is possible to find atypical features in MODY, like positive autoantibodies. In fact, one study found that 17% of MODY individuals and 34% of T2DM individuals were positive for at least one β-cell autoantibody (10).

T2DM poses yet another problem for MODY diagnosis as its increasing prevalence in youths complicates the distinction of MODY from early onset T2DM by hampering the age of onset and family history criteria's usefulness. Also, MODY patients are typically lean, as opposed to early onset T2DM individuals. However, the increasing prevalence of adolescent and young adult obesity means more MODY patients present with increased body mass index (BMI), which could confound the diagnostic process (5,8).

Thus, genetic testing is essential for a correct diagnosis of MODY and its subtype. Downstream advantages include: 1) implementation of optimal treatment, which leads to improved quality of life and improved glycemic control, if an individual transfers from insulin therapy to oral treatment with

sulphonylurea (6,19); 2) estimation of patient prognosis; 3) prompt screening for other abnormalities associated with MODY gene mutations; 4) allow screening of proband relatives to ascertain their carrier status (3,5). For an in-depth comparison between MODY, T1DM and T2DM see table 6.1 in the supplementary material section.

1.3 *GCK* and GCK-MODY

GCK, also known as hexokinase IV or hexokinase D, is a product of the *GCK* gene, located on chromosome 7 (7p13). According to NCBI, *GCK* has 13 exons, nine transcript variants and generates three tissue-specific isoforms via alternative splicing. Transcript variant 1 is expressed in pancreatic β-cells, as well as multiple other cells, and encodes isoform 1, with 465 amino acid residues. Transcript variants 2 and 3 are expressed in the liver and encode isoforms 2 and 3, the major and minor hepatic isoforms, with 466 and 464 amino acid residues, respectively. Transcript variant 1 possesses a specific first exon with an exclusive 5' UTR, generating a unique N-terminus in isoform 1. Transcript variants 2 and 3 have a liver specific first exon, distinct from the one in transcript variant 1. Transcript variant 3 has a second exon that is also liver specific, which is absent from transcript variant 2. Thus, the three isoforms are discernible through their overall distinct N-termini, being otherwise identical among them.

GCK has two distinct promoters, separated by 30 kbp of genomic DNA, that modulate tissue-specific expression. The upstream neuroendocrine promoter and its adjacent exon prompt transcription in β-cells, as well as enteroendocrine cells, like the intestinal K- (20) and L-cells (21), glucose-sensitive neurons and multiple other cells in diverse tissues. Conversely, the downstream liver promoter and its specific adjacent exon incite transcription exclusively in hepatocytes (22,23). These alternative promoters and their leader exons are at the basis of GCK isoforms' distinct N-termini.

In the liver, GCK synthesis is induced by - and solely dependent on - insulin, thereby reflecting the organism's nutritive state, either fed or fasting, with correspondingly high and very low GCK concentrations. Unlike insulin, glucagon has the opposite effect, suppressing *GCK* expression. However, in β-cells, *GCK* is constitutively expressed, regardless of the body's nutritive state and, by extension, insulin levels (22).

GCK is a monomeric protein with a small and large domain separated by a deep cleft, where the glucose binding site is located. This cleft is composed from small domain residues (Thr168 and Lys169), large domain residues (Glu256 and Glu290) and connecting region II residues (Asn204 and Asp205). GCK is allosterically regulated and seems to have three structural conformations and two catalytic cycles. Two conformations have been solved: a catalytically inactive super-open conformation, with low glucose affinity; and a catalytically active closed conformation, with high glucose affinity. Evidence suggests an intermediary active open conformation, with high glucose affinity, free of substrates and other products. The transition between the super-open conformation, where the glucose binding site is exposed to the solvent, and the open conformation corresponds to a slow catalytic cycle, due to its complex and time consuming molecular rearrangements. In contrast, due to its less complex reorganization, transition between open and closed conformations is fast and easily reversible, corresponding to a fast catalytic cycle (16,22,24).

GCK plays an important role in glucose homeostasis. Due to its low affinity towards glucose and cooperative kinetics, it acts as a highly sensitive glucose sensor in cells (25). Because it is not inhibited by its product, GCK is able to continuously carry its activity (23). This enzyme catalyzes transfer of a phosphate from ATP to glucose, the first and rate limiting step of glucose metabolism, to

generate glucose-6-phosphate (G6P). GCK activity peaks with hyperglycemia, being directly proportional to ambient glucose concentration (4,16). In hepatocytes, at low glucose concentrations, GCK regulatory protein (GKRP) binds and allosterically inhibits GCK in the nucleus, which adopts the super-open conformation. A raise in glucose concentration favors GCK release from GKRP, allowing GCK-glucose binding, accompanied by conformational change to the closed form and its translocation to the cytoplasm, where it can catalyze glucose phosphorylation (16,22,25). G6P's increasing concentration via GCK action leads to an increase in glycogenesis, as well as glycolysis and glucose oxidation, contributing to postprandial glucose level regulation (22).

In the β-cell, GCK facilitates insulin release. G6P generated by GCK undergoes glycolysis to produce pyruvate, which then enters the tricarboxylic acid (TCA) cycle to yield ATP. This raises the ATP/ADP ratio, closing ATP-sensitive potassium channels and depolarizing the plasma membrane. In turn, voltage-gated Ca^{2+} channels open, resulting in Ca^{2+} influx and, ultimately, insulin exocytosis. Like in hepatocytes, GCK controls the phosphorylation step, the first in glucose metabolism. Initiation of this pathway results in glycolysis and glucose oxidation, effectively meaning the initial reaction, catalyzed by GCK, exerts a great amount of control over glucose's metabolic flux (22,25).

Given GCK's role in glucose metabolism and insulin release, it is no surprise that *GCK* mutations can result in both hyper- and hypoglycemia (16). Heterozygous loss-of-function mutations lead to a decreased phosphorylation rate. In the liver, this decreases the amount of glycogen synthesized, hindering postprandial glucose regulation (26). In β-cells, insulin secretion regulation is impaired, as evidenced by a shift to the right in the glucose insulin secretion rate dose-response curve, an expression of a decrease in these cells' responsiveness to glucose (27). Although still under tight homeostatic control, this means the glycemic threshold for insulin release is set at a slightly higher concentration, when compared to healthy controls (4). The overall net result of these mutations is hyperglycemia.

Although rare, homozygous loss-of-function mutations also lead to hyperglycemia. These mutations bring about total GCK deficiency that manifests itself in the form of permanent insulin-requiring diabetes with neonatal onset. This more severe phenotype can also result from compound heterozygous loss-of-function mutations and is known as permanent neonatal diabetes mellitus (PNDM) (16). In opposition, heterozygous gain-of-function mutations, which have been increasingly reported, shift GCK to the active conformation at lower glucose concentrations, resulting in hyperinsulinemic hypoglycemia (4,16,28).

Over 600 different mutations have been identified scattered throughout the *GCK* gene (16). Remarkably, the heterozygous loss-of-function mutations all result in an identical phenotype, due to the compensatory action of the remaining wild type allele (4). This phenotype is known as GCK-MODY or MODY2 and is characterized by asymptomatic mild stable hyperglycemia present from birth. Being asymptomatic and non-progressive, hyperglycemia often remains undetected and individuals are usually incidentally diagnosed during routine investigations. The same applies to these individuals' relatives, who might be unaware of their carrier status and remain undiagnosed or misdiagnosed. Often, one parent has mild hyperglycemia and family history of T2DM or gestational diabetes (GDM) is common (4,5).

Notably, GCK-MODY individuals retain good homeostatic control over blood glucose, as evidenced by glucose levels' small increase at the 120 minute mark of an oral glucose tolerance test (OGTT). In fact, one study found that, at this mark, 95% of individuals display glucose levels below 83 mg/dL (4.61 mmol/L) (4,5). This, coupled with hyperglycemia's non-progressive nature, makes micro- and

macrovascular complications rare (4). Additionally, treatment with oral hypoglycemic agents (OHA) or insulin therapy may be ineffective in decreasing glycated hemoglobin (HbA1c), as administration of low dose exogenous insulin leads to a compensatory decrease in endogenous insulin secretion, meaning glucose remains unaltered and HbA1c values are near normal (3,4,29).

In pregnant women, insulin therapy may be necessary to prevent consequences for the baby (6). If the baby does not inherit a heterozygous loss-of-function *GCK* mutation from its afflicted mother, it will be at risk of macrosomia, due to increased insulin secretion and insulin-stimulated growth secondary to maternal hyperglycemia. However, if the baby inherits the mutation from the father and the mother is unaffected, there will not be enough glucose to stimulate the appropriate insulin secretion for normal fetal growth and the baby will be born underweight. If both mother and baby carry mutations, the baby will have the necessary glucose to stimulate the proper insulin secretion for normal fetal growth and, thus, the mutations cancel each other and the baby is born with normal weight (3–5,30).

In lean young adults, teens and children, the following features are usually indicative of GCK-MODY (see table 6.1 for additional information) (3,4):

- Persistent fasting hyperglycemia on at least three separate occasions spanning months to years (fasting plasma glucose (FPG) 99-153 mg/dL or 5.5-8.5 mmol/L);
- HbA1c near normal (typically under 7.5%);
- OGTT increment (120 minute glucose – 0 minute glucose) under 54 mg/dL or 3.0 mmol/L;
- Persistent fasting C-peptide production (stimulated serum C-peptide>200 ρmol/L);
- Negative pancreatic autoantibodies;
- One parent will generally have mild hyperglycemia (FPG 99-153 mg/dL or 5.5-8.5 mmol/L) unless in the presence of a *de novo* mutation.

For pregnant women, the diagnostic criteria are (3):

- Persistent fasting hyperglycemia before, during and after pregnancy (FPG 99-144 mg/dL or 5.5-8.0 mmol/L);
- At least one OGTT increment under 82.8 mg/dL or 4.6 mmol/L during or after pregnancy;
- Absence of family history should not exclude GCK-MODY as one parent may have mild diabetes that remains undetected.

1.4 *HNF1A* and HNF1A-MODY

HNF1A is a transcription factor encoded by the *HNF1A* gene on chromosome 12 (12q24.31). *HNF1A* has 10 exons and generates multiple transcripts via alternative splicing, which encode three different isoforms (figure 1.1) (31,32). According to Uniprot, isoform A encodes the longest protein of the three, with 631 amino acid residues, while isoforms B and C encode shorter truncated versions. Isoform B has 572 amino acid residues and isoform C has 524 amino acid residues. All three isoforms are present in the liver, kidney, pancreas, isolated islets and intestines, but their expression patterns differ. Isoform A is predominant in the liver, kidney and fetal pancreas, while isoform B is the most abundant in the mature pancreas, where it seems to play a role in the continued maintenance of β-cell function. Thus, *HNF1A* expression differs in space and time, according not only to tissue but also with developmental stage. This differential expression pattern may reflect differences in function between isoforms, as isoform A is thought to have the lowest transactivation potential of the three, or reflect differential activation of downstream effectors (33).

HNF1A has three functional domains: a dimerization domain, a DNA-binding domain and a transactivation domain (figure 1.1). Located in the nucleus, HNF1A acts as a homodimer, the predominant form in liver and pancreas, and plays a regulatory role in the expression of several genes in multiple tissues during embryonic development. In the liver, HNF1A binds several liver-specific genes like albumin or fibrinogen's alpha and beta chains, and a total of at least 222 genes in hepatocytes (34). HNF1A also plays a role in regulating insulin expression (also binding the insulin receptor promoter) and in the development, proliferation and cell death in the mature β-cell. Unlike in GCK-MODY, *HNF1A* mutations do not seem to influence birth weight, as *in utero* β-cell function is normal, which also supports *HNF1A*'s differential expression throughout an individual's life (4,5,34).

HNF1A shows great allelic heterogeneity (31), with over 400 pathogenic variants reported, most of them in exons 2 and 4. One study found that missense mutations dominate, corresponding to 54.7% of reported mutations, followed by frameshifts at 21.7%, splice site mutations at 8.7%, promoter region mutations at 1.9% and partial or whole gene deletions at 1.2% (5,15). Another study found that most dimerization and DNA-binding domain mutations were missense (74%); and truncating mutations were predominant in the transactivation domain (62%). Some transactivation domain missense mutations may result in just a mild phenotype or not even result in diabetes at all. Truncating mutations tend to lead to similar phenotypes, most likely via nonsense-mediated decay (NMD), ultimately resulting in haploinsufficiency (5,15,31).

Heterozygous *HNF1A* mutations cause HNF1A-MODY. These mutations are highly penetrant, with 63% of carriers developing diabetes by 25 years of age and 96% by 55 years, but tend to display variable expressivity, as some individuals are normoglycemic while others are hyperglycemic. This could point to modifier gene involvement (4,31). As *HNF1A* is expressed in several tissues and directly or indirectly influences multiple gene expression, *HNF1A* mutations have a wide range of possible consequences. In humans, glucose- and amino acid-induced insulin secretion is impaired, which has been associated with defective islet-cell glycolytic flux and oxidative phosphorylation. To shed some light on the underlying processes of HNF1A-MODY, one study used knockout mouse pancreatic islets and hepatocytes and found evidence that *HNF1A* deficiency alters gene expression for proteins involved in glucose transport, like glucose transporter 2 (GLUT2), and glucose metabolism, like key mitochondrial enzymes. Additionally, the same study found that numerous genes with functions spanning from glycolysis to gluconeogenesis to amino acid derivation to the TCA cycle were downregulated. As glucose metabolism is affected, glucose-dependent gene expression in *HNF1A* deficient islets is also affected. β-cell proliferation was also demonstrated to be impaired (35). However, it should be noted that *HNF* gene expression profiles differ between humans, mouse and rat, especially for *HNF1A* (32), meaning conclusions from animal models should be carefully interpreted.

HNF1A mutation location has been shown to influence age of onset, with mutations in exons 1-6 associated with a lower age at onset than mutations in exons 8-10 (3). This happens because mutations in exons 1-6 affect isoforms A, B and C, having a potentially more severe phenotype, while mutations in exons 8-10 only affect isoform A, which is not the predominant isoform in the mature pancreas. One study found this genotype-phenotype relationship to only be true for missense mutations and unrelated to the protein's affected functional domain (33). However, another study found that patients with truncating mutations were, on average, diagnosed four years earlier than patients with missense mutations, reflecting the potentially more severe consequences of truncating mutations. The same study also found that individuals with dimerization and DNA-binding domain missense mutations were 10 years younger at onset than individuals with transactivation domain missense mutations. This could be due to whether one, two or three isoforms are affected or due to a particular position's importance within a domain in relation to its function (figure 1.1) (3,4,31).

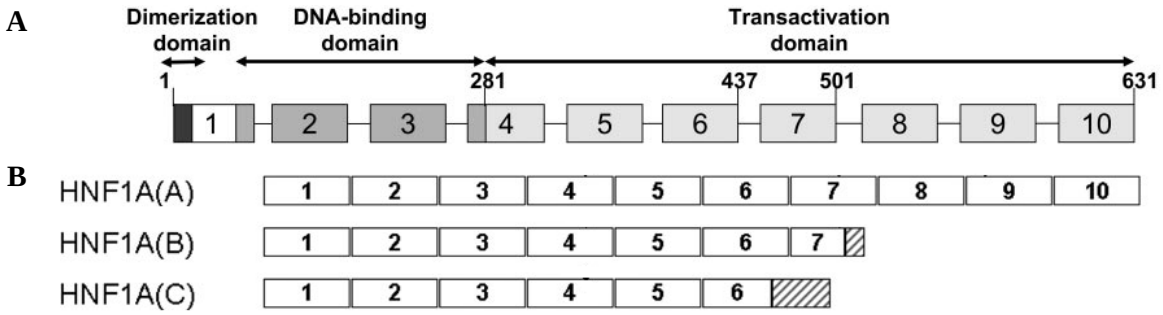


Figure 1.1: *HNF1A* structure. Gene structure (A) and isoforms A, B and C structure (B). Numbered boxes correspond to exons and hatched boxes to incorporated intronic sequences. Numbers 1 to 631 correspond to amino acid residue positions. Adapted from (31,33).

HNF1A-MODY is characterized by a progressive insulin secretory defect that leads to a more severe hyperglycemia than GCK-MODY. Interestingly, these individuals often present normal FPG, which might result from sufficient insulin secretory capacity in HNF1A-MODY's early stages coupled with relative insulin sensitivity and low BMI. As a result, normal glucose concentrations are achieved. In fact, it has been demonstrated that non-diabetic mutation carriers have normal insulin secretion at normal glucose concentrations (36). It has also been demonstrated that non-diabetic mutation carriers already present β-cell deficiency before onset of hyperglycemia, which could be delayed by normal or increased insulin sensitivity. This means that, in the early stage of diabetes, these individuals do not require exogenous insulin, as can happen later (4,5,37). Regardless, HNF1A-MODY's progressive nature has time on its side as hyperglycemia worsens and patients ultimately tend to display poor and deteriorating glycemic control, with vascular complications – micro and macro – ensuing as frequently as they do in T1DM and T2DM. This is supported by large OGTT increments and impaired insulin secretion rate with rising glucose concentrations (4,36).

In addition to the aforementioned decreased β-cell proliferation, increased cell death through apoptosis has been proposed to explain the declining β-cell function behind HNF1A-MODY's progressive profile (4,5). Despite this, patients present normal glucose levels at birth and are generally lean, with diabetes onset typically occurring in adolescence or early adulthood, before the age of 25 years (4). These individuals are more likely to progress from a state of normal glucose tolerance to impaired glucose tolerance (IGT) and, finally, to diabetes. All of this is to be expected given HNF1A's regulatory actions in the mature β-cell.

HNF1A is expressed outside the pancreas and these patients often present extra-pancreatic features. HNF1A plays a role in glucose reabsorption in the proximal renal tubules, by directly regulating sodium glucose cotransporter 2 (*SGLT-2*) expression (37). Loss-of-function *HNF1A* mutations hamper this process, resulting in decreased *SGLT*-2 expression and reduced renal glucose reabsorption, ultimately leading to glycosuria that manifests itself even before onset of hyperglycemia. In fact, 38% of non-diabetic mutations carriers were found to have glycosuria subsequent to an OGTT. In the early stage of HNF1A-MODY, OGTT will show a marked increase in glucose concentration (3). In comparison with GCK-MODY, these individuals present higher OGTT increments, with lower insulin concentrations throughout the test, illustrating their insulin secretion defect (4,36,37).

Additionally, HNF1A-MODY patients tend to have elevated HDL, unlike in T1DM and T2DM, where HDL concentrations tend to be normal and low, respectively. Despite this, cardiovascular risk does not seem to decrease, as coronary heart disease incidence sits between T1DM and T2DM (4). Interestingly, in rare cases, liver adenomatosis occurs due to upregulated expression of proliferation and cell cycle regulatory genes secondary to somatic *HNF1A* biallelic loss-of-function (31,35).

In typically lean diabetic individuals with onset before the age of 25 years and a family history of diabetes, the following features are normally indicative of HNF1A-MODY (3,4):

- OGTT increment (120 minute glucose – 0 minute glucose) over 63 mg/dL or 3.5 mmol/L;
- Absence of pancreatic autoantibodies;
- Persistent fasting C-peptide production (stimulated serum C-peptide >200 ρmol/L);
- Glycosuria at glucose levels under 180 mg/dL or 10 mmol/L;
- Normal or elevated HDL levels (>1.3 mmol/L);
- Marked sensitivity to the OHA sulphonylurea (with improving glycemic control).

OGTT will have higher diagnostic rate for HNF1A-MODY than GCK-MODY, due to its often normal FPG and greater OGTT increment, reflecting its insulin secretory defect and resulting impaired homeostasis capacity. In contrast, GCK-MODY is more easily diagnosed with FPG alone, as fasting hyperglycemia is persistent and OGTT increment is lower, due to relatively conserved insulin response that translates into a more capable homeostatic capability (36). Naturally, the combined use of diversified diagnostic criteria yields the highest diagnostic rates.

1.5 *HNF4A* and HNF4A-MODY

HNF4A is a transcription factor encoded by the *HNF4A* gene, located in chromosome 20 (20q13.12). *HNF4A* has 13 exons, with four variants for exon 1 (1A, 1B, 1C and 1D) (figure 1.2), and two promoters. The P2 (pancreatic) promoter is 46 kb upstream of the P1 (hepatic) promoter, controlling expression in the pancreas, though it is also used in hepatocytes (38,39), and its transcripts contain exon 1D (40). The P1 promoter seems to drive expression in the liver, kidney and fetal pancreas, but not in the mature pancreas (39), and its transcripts contain exon 1A (40). However, there are conflicting reports, with one study reporting P1-driven expression in the pancreas (40). According to NCBI, *HNF4A* generates multiple transcripts to encode 10 different isoforms, 1 through 10. These isoforms differ either at their N-terminus, C-terminus or both, by using alternate up- or downstream start codons and alternate splice and polyadenylation sites. Isoforms 1-6 are expressed via the hepatic promoter, while isoforms 7-9 are expressed via the pancreatic promoter, with isoform 8 being the most expressed in adult pancreas (39). Isoforms 2 and 8 are the alternatively spliced versions of isoforms 1 and 7, respectively (40). So, much like *HNF1A*, *HNF4A* is differentially expressed in space and time. Globally, *HNF4A* is expressed in the liver, where it is the most abundant DNA-binding protein; kidney; pancreas, where it controls approximately 11% of islet genes (41); and intestines. In the mature β-cell, *HNF4A* expression is negatively regulated by itself (39).

Like HNF1A, HNF4A is located in the nucleus and acts as a homodimer to bind DNA. HNF4A is part of the same complex transcription factor network that includes HNF1A and modulates gene expression in multiple tissues during embryonic development, being undoubtedly important as knockout animals are not viable (4,33). However, it seems HNF4A acts in a much larger number of hepatocyte and β-cell genes than HNF1A, by directly binding almost half of actively transcribed genes (34). One of the genes regulated by *HNF4A* is none other than *HNF1A*, which in turn also regulates *HNF4A* expression. As this regulatory network is linked to several nutrient transport and metabolic pathways, and considering the *HNF4A*-*HNF1A* regulatory loop, a disruption in these or other points in the network can result in a pleiotropic effect due to inefficient execution of cellular genetic programming, ultimately reflected in cellular dysfunction (35). In the liver, *HNF4A* is crucial for hepatocyte morphological and functional differentiation, glycogen storage and generation of hepatic epithelium (42), while also regulating expression of genes involved in gluconeogenesis and lipid

metabolism (41). *HNF4A* also plays a role in pancreatic development, β-cell differentiation and function maintenance (39), and insulin secretion, by directly activating the *INS* promoter (40).

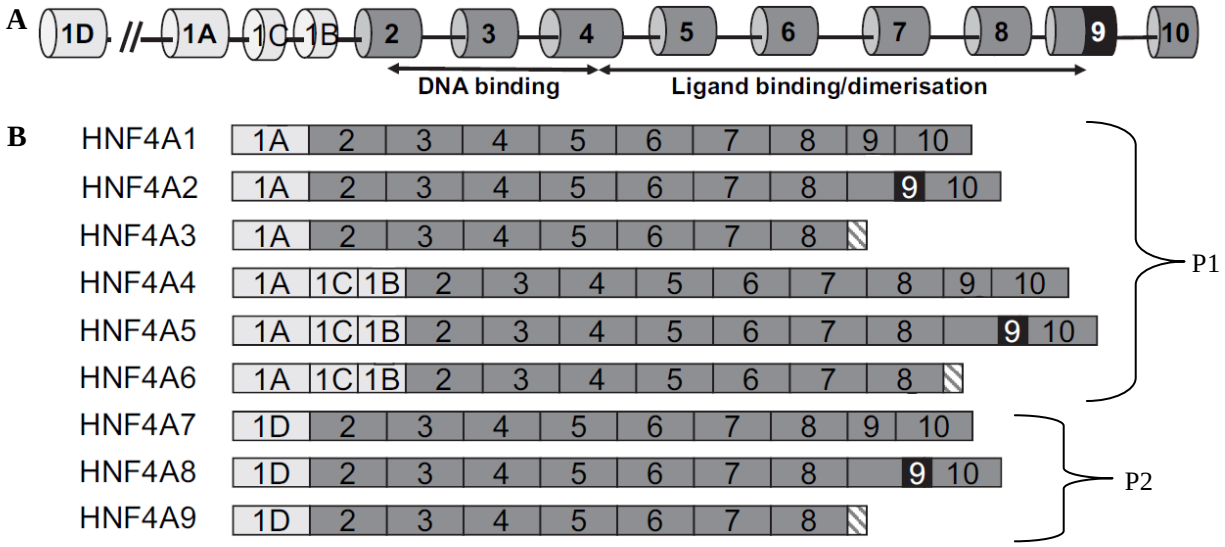


Figure 1.2: *HNF4A* gene and isoform structures. A: alternate first exons (1A-1D) in light grey and remaining exons in dark grey. B: HNF4A isoforms' structural composition. Hatched boxes correspond to retained intronic sequences. In black, a 10 amino acid residue insertion in exon 9. P1 - hepatic promoter; P2 - pancreatic promoter. Adapted from (39).

Heterozygous *HNF4A* mutations cause HNF4A-MODY, which accounts for 3-10% of MODY cases, depending on the sources (4,5). These mutations can be located in the gene itself or its P2 promoter (39). HNF4A-MODY is characterized by a phenotype identical to that of HNF1A-MODY, with progressive β-cell dysfunction. However, *HNF4A* mutations have lower, yet variable, penetrance, as some individuals remain free of diabetes in their fourth decade (3,4). Penetrance varies with affected isoforms, as mutations affecting all isoforms show the highest penetrance (39). Over 100 mutations have been reported in *HNF4A*: 58.3% are missense mutations; 11.7% are frameshifts; 9.7% are nonsense mutations; 5.8% are splice site mutations; 5.8% are promoter region mutations; and 1.9% are partial and whole gene deletions. Although mutations are found throughout *HNF4A*, most of them are located in exons 7 and 8 (15).

Unsurprisingly, HNF4A-MODY's progressive β-cell defect translates in failure to appropriately increase insulin secretion in response to rising glucose concentrations. Like in HNF1A-MODY, mutation position is correlated with age at diagnosis, which could be explained by the differential expression patterns of *HNF4A*'s isoforms. Mutations in exons 2-8, which affect all isoforms, are associated with earlier diagnosis, as opposed to mutations in exons 9 and 10, which affect isoforms 1, 2, 4, 5, 7 and 8, and are associated with later diagnosis. Mutations affecting P2-derived isoforms are associated with a later age at diagnosis. Mutation type and location within isoform structure do not seem to influence age at diagnosis (39).

As phenotypes are very similar, *HNF4A* mutations should be screened for when clinical features are highly suggestive of HNF1A-MODY but no *HNF1A* mutations are found (3). However, unlike HNF1A-MODY, *HNF4A* mutations are frequently associated with fetal macrosomia, as offspring of an affected mother or father have 50-56% risk of being born with excess weight (3,4). Increased *in utero* insulin secretion is associated with increased birth weight. In the neonatal period, increased insulin secretion can lead to hyperinsulinemic hypoglycemia, presenting in the first week of life. This condition affects approximately 10% of mutation carriers and can be transient or prolonged, ultimately transitioning to diabetes, though it is not known how nor when (4,5,39).

Due to its multi-tissue expression pattern and wide action range, *HNF4A* mutations also have extra-pancreatic effects like low HDL, low apolipoprotein levels (apo A1, apo A2[1], apo CIII, apo B), low triglycerides and elevated LDL. Unlike HNF1A-MODY, HNF4A-MODY individuals do not present glycosuria. Thus, the following features are indicative of HNF4A-MODY (4,5):

- OGTT increment (120 minute glucose – 0 minute glucose) over 63 mg/dL or 3.5 mmol/L;
- Persistent fasting C-peptide production (stimulated serum C-peptide >200 ρmol/L);
- Negative pancreatic autoantibodies;
- Macrosomia and/or neonatal hyperinsulinemic hypoglycemia;
- Low HDL, low triglycerides, elevated LDL and sensitivity to the OHA sulphonylurea.

1.6 *HNF1B* and HNF1B-MODY

HNF1B is a transcription factor encoded by the *HNF1B* gene, located in chromosome 17 (17q12). According to NCBI, *HNF1B* has 11 exons and generates three transcripts via alternative splicing to encode three isoforms. Isoform 1 is the longest, with isoforms 2 and 3 lacking an internal segment and differing in their C-terminus, respectively. Like HNF1A, HNF1B has a dimerization domain, a DNA-binding domain and a transactivation domain (figure 1.3) (44). *HNF1B* is expressed in early stage embryonic development in the thymus, lung, kidney, liver, pancreas, bile ducts, stomach, intestines and genital tract, being vital for embryonic survival (4,44).

Also located in the nucleus, HNF1B acts as a homodimer or heterodimer. When acting as a heterodimer, HNF1B is coupled with HNF1A, a likely reflection of their homology, which includes the homeodomain and dimerization domain (44). *HNF1B* and *HNF1A* are actually paralogs with known interchangeable functions in several contexts, with HNF1B occupying a subset of direct HNF1A target genes to assure normal expression (35). HNF1B also binds *HNF4A*'s P2 promoter, adding another nexus in this transcription factor network that regulates embryonic gene expression (39). *HNF1B* plays a role in nephron development, regulates embryonic pancreatic development and regulates expression of several genes involved in cholesterol and sphingolipid metabolism (45). One study assigns *HNF1B* a role in hepatic insulin sensitivity control (46). Another study suggests HNF1B functions as a classic transcriptional activator and as a bookmarking factor that marks target genes for rapid transcriptional reactivation after mitotic silencing via chromatin condensation (47).

Heterozygous *HNF1B* mutations result in HNF1B-MODY, which accounts for 1-5% of all MODY cases. Penetrance is highly variable, as age at diagnosis ranges from 0 to 61 years (4,5). Mutation carriers also present variable phenotypes, with a wide range of clinical features. In one case, opposing clinical features have been associated to the same mutation within the same family (44,48). This variability is likely due to *HNF1B*'s connections within its regulatory network. Over 65 mutations have been reported and approximately 28% of individuals present with full allele deletion (4,5,49). Moreover, *de novo* mutations are frequent, comprising as much as half of cases, meaning family history may be absent (4). Loss-of-function, dominant-negative and gain-of-function mutations have all been reported, falling in the missense, nonsense, frameshift, insertion/deletion, and splice site categories (44). One study found that mutations were predominantly located throughout the DNA-binding domain, mostly in exons 2 and 4, but rarely affected exon 3, which is entirely included in this domain. The same study and one other report that intron 2's donor splice site was a hot spot for mutation (44,49). This second study also found the majority of mutations to be confined to the first four exons, which encode the dimerization and DNA-binding domains, with exon 2 being the most

[1] *HNF4A* plays a role in apo A2 transactivation (43).

affected. This region is critical for transcriptional activity, meaning mutations in this area will result in reduced target gene activation. No missense mutations were found in the transactivation domain (44).

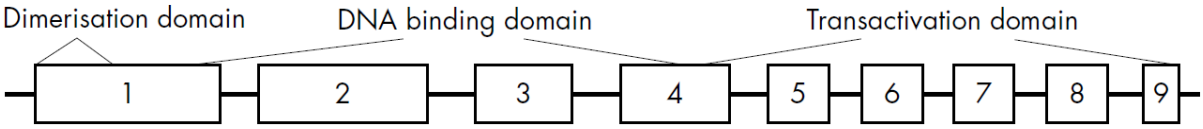


Figure 1.3: *HNF1B* structure. Genomic structure and domain design are similar to those of *HNF1A*. Exon count refers to transcript NM_000458. Adapted from (44).

Unlike the other HNF MODY, in approximately 50% of *HNF1B* mutation carriers, diabetes results from a combination of β-cell dysfunction and insulin resistance. Overall pancreatic atrophy, pancreatic dysplasia and exocrine dysfunction are common. Another distinctive feature is absence of sulphonylurea sensitivity, with insulin treatment required at an early phase. *In utero* insulin secretion is reduced, resulting in significant birth weight decrease (800-900 g). Like in HNF1A- and HNF4A-MODY, insulin secretion is compromised as glucose concentrations rise, though initially maintained at normal glucose concentrations. As such, microvascular complications are also frequent (4,5,44).

As *HNF1B* is expressed in several tissues during embryonic development, *HNF1B* mutations can also result in extra-pancreatic abnormalities. The most afflicted organ is the kidney, usually affected by renal cysts and diabetes (RCAD) syndrome, though renal structure anomalies seem to precede and be more prevalent in relation to diabetes. In fact, diabetes is not an essential feature for *HNF1B* mutation identification. Renal cysts are the most common abnormality, though renal dysplasia and renal tract malformations, like horseshoe kidney, have been reported. Moreover, less than 6% of HNF1B-MODY patients have normal renal function and about half have end-stage renal failure. Genital tract anomalies, like vaginal aplasia or azoospermia, have also been reported, but penetrance is incomplete. Other associated anomalies are liver dysfunction and abnormal liver function tests, specially alanine aminotransferase (ALT) and gamma-glutamyl transferase (GGT); gallbladder dysfunction; hyperuricemia and resulting gout; and hypomagnesemia (4,5,44,49). Thus, the following features are indicative of HNF1B-MODY (4):

- OGTT increment (120 minute glucose – 0 minute glucose) over 63 mg/dL or 3.5 mmol/L;
- Persistent fasting C-peptide production (stimulated serum C-peptide >200 ρmol/L);
- Negative pancreatic autoantibodies;
- Elevated creatinine and uric acid secondary to compromised renal function;
- Elevated liver enzymes and low magnesium.

1.7 Rare types of MODY

The aforementioned types of MODY account for the vast majority of cases. However, a few cases have different genetic etiologies. Mutations in *PDX1* (5), *NEUROD1* (4,5), *KLF11* (5,50), *CEL* (4,5,51), *PAX4* (5,52), *INS* (5), *BLK* (4,5,53), *ABCC8* and *KCNJ11* (4,5) have all been associated with MODY, each accounting for less than 1% of all cases (5). Additionally, some individuals with MODY have no identified mutations in these genes, having what is known as MODY X.

This study's objective is to perform the molecular characterization of Portuguese patients with a clinical diagnosis of MODY through PCR amplification and Sanger sequencing of *GCK* and *HNF1A*, and through Multiplex Ligation-dependent Probe Amplification (MLPA) for copy number variation (CNV) detection. This work is part of the large study "Molecular characterization of MODY patients" and is, to our knowledge, the first of its kind in Portugal.

2. Materials and Methods

2.1 Inclusion criteria

Between September 2013 and December 2015, subjects who presented with clinical diagnosis of MODY and met the selection criteria provided in the study protocol and clinical questionnaires were enrolled. These criteria were based on best practice guidelines defined by the European Molecular Genetics Quality Network (3). All subjects had been diagnosed with diabetes at some point in their lives. Proband candidates were typically lean, with onset of diabetes before 25 years of age, a family history of diabetes consistent with a dominant inheritance pattern and no pancreatic autoantibodies. When possible, relatives were also enrolled.

The questionnaires sought to collect data on the following parameters: sex; ethnicity; age; age at diagnosis; family history of diabetes and renal disease; presence of retinopathy, coronary disease, neuropathy and nephropathy; birth weight; BMI; FPG; OGTT; HbA1c; pancreatic autoantibodies; and treatment type (diet, OHA or insulin).

2.2 DNA extraction

Subject DNA was extracted from whole blood in ethylenediamine tetraacetic acid (EDTA) using the salting-out method, based on Lahiri and Nurnberger's publication (54). Samples were stored at 4ºC until further use.

2.3 DNA quantification and quality screen

Sample DNA concentration was assessed with NanoDrop 1000 Spectrophotometer V3.7. To ensure reproducible results, each sample was measured at least twice. To ensure minimum DNA quality, all samples ran on a 1% agarose gel (1.0 g agarose in 100 mL 1X TBE buffer) for 45 minutes at 90 volt. Samples were loaded with a mix consisting of 4.5 µL loading buffer (bromophenol blue + glycerol) and 4.5 µL bidestilled water. Sample volume was 1 µL.

2.4 Primer design, checks and preparation

Primers were designed on Primer3Plus (55), with defined target regions ideally encompassing 70 (minimum 50) or more base pairs up- and downstream the coding regions. For promoter regions, forward primers were located a few hundred base pairs upstream the first exon (494 bp in *GCK* and 469 bp in *HNF1A*), with the reverse primers ending a little over 100 bp into the corresponding first exon (104 bp in *GCK*; 117 bp in *HNF1A*). Thus, primers to amplify exonic and adjacent intronic regions, and promoter regions were designed based on reference sequences for *GCK* (NG_008847.1; NM_000162.3) and *HNF1A* (NG_011731.2; NM_000545.6). All primer binding sites were checked for SNP's using the online tool SNPCheck. At the time of design, all primer binding sites were SNP free. See supplementary material for primer sequences.

In silico PCR was performed online on University of California Santa Cruz (UCSC) for all primer pairs. Predictions showed all primer pairs amplified the intended target regions. The online UCSC BLAST-like Alignment Tool (BLAT) was used on all defined PCR products (target and included regions) to ascertain sequence similarity against genomic target regions and ensure the predicted products were the ones intended. Manual double-checks on the genomic RefSeqs were also performed for the same purpose.

PCR primers were synthesized by Frilabo. Dehydrated primers were stored at 4°C until rehydration. All primers were rehydrated with sterile bidestilled water in laminar flow cabinets, to 100 µM (100 ρmol/µL) stock solutions. Working solutions were made from stock solutions at 10 µM (10 ρmol/µL) for PCR (10 µL from primer stock solution + 90 µL sterile bidestilled water) in 1.5 mL Eppendorf tubes® and, from the latter, working solutions were made at 2 µM (2 ρmol/µL) for Sanger sequencing (20 µL from PCR working solution + 80 µL sterile bidestilled water) in 0.2 mL tubes. All stock and working primer solutions were stored at -20°C.

2.5 PCR optimization

All PCR reactions were optimized as follows: triplicates from two healthy samples were initially amplified for each fragment at 59°C, 60°C and 61°C annealing temperatures. If necessary, annealing temperature was adjusted 1ºC (up or down) until optimum or satisfactory amplification was achieved. After running agarose gels, optimum annealing temperatures were set for each fragment.

2.6 PCR reactions

All PCR reactions, for both optimization and fragment amplification, were performed on either a Biometra T3000 thermocycler, an Applied Biosystems 2720 Thermal Cycler or a GeneAmp® PCR System 9700. Bioline reagents were used (dNTP's, 10x NH4-based reaction buffer, 50 mM $MgCl_2$ solution, and BIOTAQ™ DNA Polymerase) and kept stored at -20ºC. Standard PCR master mix and cycling program are available in the supplementary material. When possible, DNA samples were used at 100 ng/µL. When their concentration was below 100 ng/µL, samples were used as they were.

2.7 Agarose gel electrophoresis

PCR products were subjected to electrophoresis in 1.5% gel, polymerized from 1.5 g SeaKem® LE Agarose in 100 mL 1X TBE buffer (diluted from UltraPure™ 10X TBE Buffer), and 2.0 µL SYBR® Safe DNA Gel Stain. Gel was submerged in 1X TBE buffer and wells were loaded with previously mixed 5.0 µL loading buffer (bromophenol blue + glycerol) and 5.0 µL PCR product. Thermo Scientific GeneRuler 100 bp DNA Ladder (SM0241), 100-1000 bp, was used to check amplicon size. Typically, each run was 50 minutes at 90 volt. Electrophoresis were performed in Bio-Rad submerged horizontal electrophoresis cells (Mini-Sub® Cell GT Cell, Wide Mini-Sub Cell GT Cell, Sub-Cell® GT Cell and Sub-Cell® Model 192 Cell) with Bio-Rad power supplies (PowerPac™ HC or PowerPac™ Basic). Agarose gels were analyzed under UV light, with Safe Imager™.

2.8 PCR product purification and Sanger sequencing reaction

PCR product purification was performed with ExoSAP-IT™ PCR Product Cleanup Reagent, following manufacturer protocol, on the same thermocyclers as PCR reactions.

Purified PCR products were then sequenced using the BigDye® Terminator v1.1 Cycle Sequencing kit. These reactions were also performed on the same thermocyclers as PCR reactions and then sequenced on Applied Biosystems® 3130xl Genetic Analyzer. See supplementary material for internal sequencing mix and cycling program.

2.9 Sequence analysis and variant interpretation

Chromatogram analysis was performed with the Staden Package software (Trev, Pregap4 and Gap4 programs) by aligning all chromatograms for each fragment and comparing them against each other and the correspondent RefSeq. For Pregap4 module configuration, see supplementary material.

Sequence variants were analyzed and interpreted by querying open online databases, namely NCBI's dbSNP, 1000Genomes Browser, PubMed and ClinVar; Human Gene Mutation Database (HGMD), Exome Variant Server (EVS) and Exome Aggregation Consortium (ExAC); and also by querying *in silico* predictive software such as MutationTaster (56), PolyPhen-2 (57) and PROVEAN/SIFT (58,59). Final classification followed American College of Medical Genetics and Genomics (ACMG) and Association for Molecular Pathology (AMP) recommendations (60). Variant nomenclature was checked with Mutalyzer (61).

2.10 MLPA

MLPA was performed to investigate the presence of CNV on samples with no detectable pathogenic/likely pathogenic sequence variants via Sanger sequencing. MLPA was performed on a Biometra T3000 thermocycler, following the MLPA® DNA Protocol version MDP-005, as recommended by MRC-Holland, and using the SALSA MLPA probemix P241-D2 MODY mix 1 description version 20 (lot D2-0413). Results were analyzed with the Coffalyzer.Net software and interpreted by querying PubMed, ClinVar, dbVar and HGMD. Protocol, probemix description and software are available at www.mlpa.com (62).

3. Results

3.1 Subjects under study

3.1.1 Number of subjects, sex, ethnicity, age and age at diagnosis

Study sample consisted of 39 individuals, including 24 probands and 15 relatives, spread across 24 families. Seventeen out of 39 were males (43.59%) and 22 females (56.41%). Among the proband group, there were nine males (37.5%) and 15 females (62.5%), whilst eight out of 15 relatives were male (53.33%) and seven were female (46.67%). Overall, the study sample was predominantly Caucasian (92.31%), with only three non-Caucasian individuals (7.69%), all of African descent and all in the proband group. Proband age ranged from 9 to 57, with a mean of 23.92 ± 11.24 years; whilst in the relatives group, age ranged from 7 to 79, with a mean of 46.47 ± 19.21 years. Overall mean age was 32.59 ± 18.33 years. Proband group age at diagnosis ranged from 1 to 35 years and presented a mean of 16.87 ± 8.29 years. Three individuals were diagnosed at 29, 33 and 35 years and information was not available for one proband. Among relatives, information was not available for six individuals. In this group, age at diagnosis ranged from 9 to 62 years, with a mean of 33.33 ± 18.48 years.

3.1.2 Family history of diabetes and renal disease

Only two subjects (8.33%) in the proband group presented without apparent family history of diabetes, with the remaining 22 (91.67%) presenting a positive family history of diabetes. Family history of renal disease was also screened and was present in one (4.17%) subject's family. Data was not available for six of the 24 probands (25%). The remaining 17 (70.83%) had no apparent family history of renal disease.

3.1.3 Vascular complications

Subjects were screened for retinopathy, coronary disease, neuropathy and nephropathy. Out of 24 probands, 22 (91.67%) had no retinopathy, coronary disease or neuropathy, with no data available for the remaining two. Only one individual (4.17%) had nephropathy and no information was available for one other individual. The remaining 22 were nephropathy free. Among relatives, data was scarce, available for three of the 15 subjects (20%) and none of them had vascular complications.

3.1.4 Birth weight and BMI

Gestation time ranged from 35 to 40 weeks and was overall normal, as only one individual, in the relatives group, was born before 38 weeks. Among probands, birth weight ranged from 1.90 to 4.00 kg, with a mean of 3.16 ± 0.52 kg. Fourteen out of 24 probands (58.33%) were born with normal weight, between 2.50 and 3.99 kg; two (8.33%) were born macrosomic, at 4.00 kg or more; and just one (4.17%) was born below 2.50 kg. Data was not available for seven probands (29.17%). Among relatives, data was only available for two out of 15 (13.33%) subjects, one born with 3.00 kg and the other 4.10 kg.

BMI was calculated before and at enrollment. In the proband group, before enrollment, BMI ranged from 13.3 to 34 kg/m^2, with a mean of 22.10 ± 5.57 kg/m^2. Eight subjects had normal BMI (33.33%), between 18.5 and 24.9 kg/m^2; five (20.83%) presented BMI below 18.5 kg/m^2; four (16.67%) were over 25 kg/m^2; and no data was available for the remaining seven (29.17%). At enrollment, BMI ranged from 14.4 to 30.43 kg/m^2, with a mean of 21.46 ± 3.88 kg/m^2. Fourteen probands (58.33%) had normal BMI, five had low BMI (20.83%), three had high BMI (12.5%) and the remaining two had no available data (8.33%). Among relatives, BMI data before enrollment was not available. At enrollment, data was available for 10 out of 15 subjects (66.67%). BMI ranged from 18.5 to 35.3

kg/m^2 with a mean of 24.62 ± 5.27 kg/m^2. Five individuals had normal BMI (33.33%), five had high BMI and no subjects presented low BMI.

3.1.5 FPG

FPG was also measured before and at enrollment. Among probands, before enrollment, FPG ranged from 89 to 600 mg/dL, with a mean of 191.06 ± 138.09 mg/dL. Not considering two extreme values (468 and 600 mg/dL) mean FPG is 145.33 ± 46.14 mg/dL. Five (20.83%) subjects had FPG between 100 and 125 mg/dL; 11 (45.83%) had FPG of 126 mg/dL or greater; and one (4.17%) had FPG below 100 mg/dL. Data was not available for the remaining seven of 24 probands (29.17%). At enrollment, plasma glucose ranged from 78 to 279 mg/dL, with a mean of 130.94 ± 46.40 mg/dL. Seven (29.17%) subjects had plasma glucose between 100 and 125 mg/dL, but for one individual postprandial plasma glucose was measured, instead of FPG. Seven (29.17%) subjects had plasma glucose levels of 126 mg/dL or above but, again, postprandial plasma glucose was measured for one individual. Two (8.33%) subjects had FPG under 100 mg/dL and the remaining eight (33.33%) had no available data.

Among relatives, there was no data on FPG before enrollment. At enrollment, FPG ranged from 93 to 144 mg/dL, with a mean of 118.17 ± 17.34 mg/dL. Seven (46.67%) subjects had FPG between 100 and 125 mg/dL; three (20%) presented FPG of 126 mg/dL or above; two (13.33%) had FPG under 100 mg/dL; and the remaining three (20%) had no available data.

3.1.6 OGTT

In the proband group, four out of 24 (16.67%) subjects had data on OGTT before enrollment. Two-hour glucose ranged from 176 to 320 mg/dL, with a mean of 239.5 ± 66.46 mg/dL. Two (8.33%) subjects had two-hour glucose under 200 mg/dL, and the other two over 200 mg/dL. At enrollment, OGTT data was available for one individual (4.17%), who had a two-hour glucose of 160 mg/dL. There was no data on OGTT among relatives, before or at enrollment.

3.1.7 HbA1c

In the proband group, before enrollment, HbA1c ranged from 4.9% to 14.1%, with a mean of 7.56 ± 2.38%. Twelve (50%) subjects had HbA1c of 6.5% or above, with seven (29.17%) between 6.5 and 7.5%. Five (20.83%) subjects had HbA1c under 6.5% and the remaining seven (29.17%) had no data. At enrollment, HbA1c ranged from 5.1 to 12.4%, with a mean of 6.62 ± 1.87. Eight (33.33%) subjects had HbA1c of 6.5% or above, with five (20.83%) between 6.5 and 7.5%. Twelve (50%) subjects had HbA1c under 6.5% and the remaining four (16.67%) had no data.

Among relatives, no data was accessible before enrollment. At enrollment, data was available for eight of 15 subjects (53.33%). HbA1c ranged from 5.5 to 7.2%, with a mean of 6.33 ± 0.58%. Three (20%) subjects had HbA1c between 6.5% and 7.5%. The other five (33.33%) presented values under 6.5%.

3.1.8 Pancreatic autoantibodies

Subjects were screened for glutamic acid decarboxylase antibodies (GADA) and islet cell antibodies (ICA) before enrollment. Data on pancreatic autoantibodies was available for 19 of 24 probands (79.17%), all negative for GADA and ICA. No information was available among relatives.

3.1.9 Treatment type (diet, OHA and insulin)

Data on diet therapy was available on 22 of 24 probands (91.67%). Nineteen (79.17%) probands were on diet and three (12.5%) were not. Among relatives, data was available for eight of 15 (53.33%) subjects. Two (13.33%) were on diet and six (40%) were not.

Data on OHA, among probands, was available on 21 of 24 (87.5%) subjects. Three (12.5%) were on OHA and 18 (75%) were not. In the relatives group, information was available on nine of 15 (60%) subjects. Seven (46.67%) were on OHA and two (13.33%) were not.

Data on insulin therapy, among probands, was available for 23 of 24 (95.83%) subjects. Thirteen (54.17%) were on insulin and 10 (41.67%) were not. In the relatives group, data was available on nine of 15 (60%) subjects. Only two (13.33%) individuals were on insulin therapy.

One (4.17%) proband and two (13.33%) relatives were on both diet and OHA. No proband or relative was on both OHA and insulin. Eleven (45.83%) probands were on both diet and insulin, but no subject among relatives was in the same situation. Only one (4.17%) proband was on all three treatment modalities at the same time, with no relative in the same situation. Two (8.33%) probands and one (6.67%) relative were not on any of the three treatment types.

3.2 Molecular analysis: Sanger sequencing and MLPA

We detected a total of 46 sequence variants via Sanger sequencing, six of them pathogenic or likely pathogenic (figure 3.1). Among *GCK* variants, 21.05% were pathogenic or likely pathogenic, versus just 7.41% in *HNF1A*. Two other sequence variants were detected by MLPA in *GCK* and *HNF1B*.

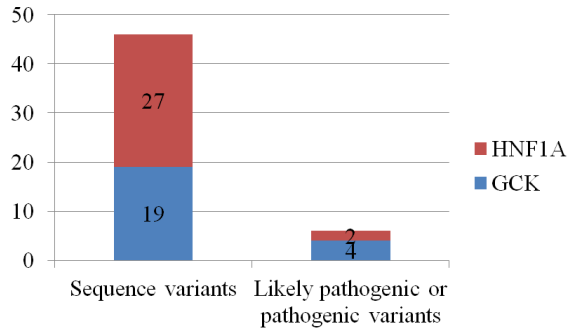


Figure 3.1: Detected sequence variants. Number of *GCK* and *HNF1A* sequence variants found via Sanger sequencing.

3.2.1 Sanger sequencing analysis: *GCK*

We detected 19 sequence variants in 19 subjects, 14 probands and five relatives. Intron 8 was the most affected region, with five variants. No variants were found in exons 2, 3, 5, 6, 8 and 9 (figure 3.2).

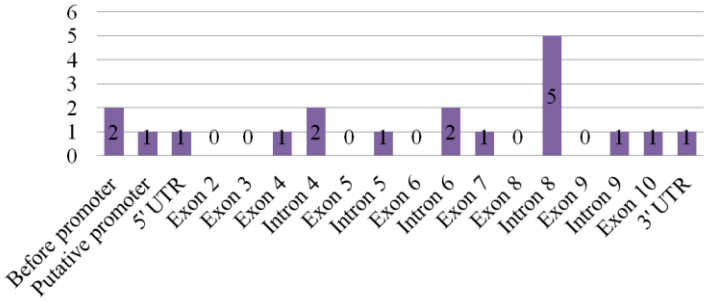


Figure 3.2: *GCK* sequence variants location. Distribution of *GCK* sequence variants throughout its genomic structure.

Unsurprisingly, most variants were located in non-coding regions, with just three out of 19 (15.79%) in coding regions. The two most common variants, c.483+87A>C and c.679+38T>C, both present in 11 probands, were located in introns 4 and 6, respectively. Conversely, all three coding region variants affected one proband each, as does a 33 bp deletion in the exon 5/intron 5 junction (figure 3.3).

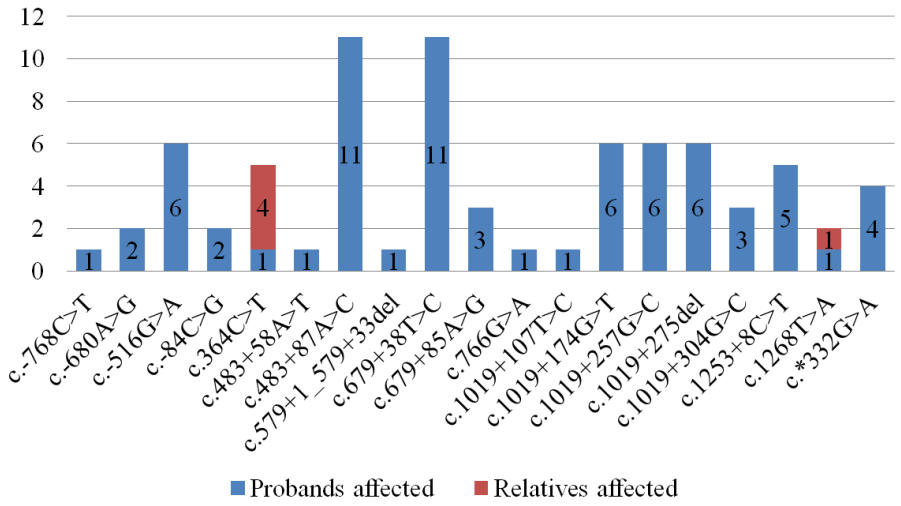


Figure 3.3: *GCK* sequence variants and affected subjects. Number of subjects affected by each *GCK* sequence variant.

Four of 19 variants can be classified as pathogenic or likely pathogenic for GCK-MODY. The first, a missense variant, is located in codon 122's first base (CTC>TTC), exon 4's first nucleotide. This c.364C>T (figure 3.4) variant is predicted to result in wild type hydrophobic leucine substitution by aromatic phenylalanine (p.(Leu122Phe)) and can be classified as likely pathogenic, as the nucleotide change occurs at a conserved position and is predicted to be deleterious by MutationTaster (0.999999982103552), PolyPhen-2 (1.000; 0.999) and PROVEAN/SIFT (-3.93/0.000). Additionally, c.364C>T cosegregates with diabetes, which is present in the proband's (III-2: 13-012) family (figure 3.4), and is associated with MODY in HGMD (CM096822) and literature (16). To our knowledge, there is no population frequency data available for this variant.

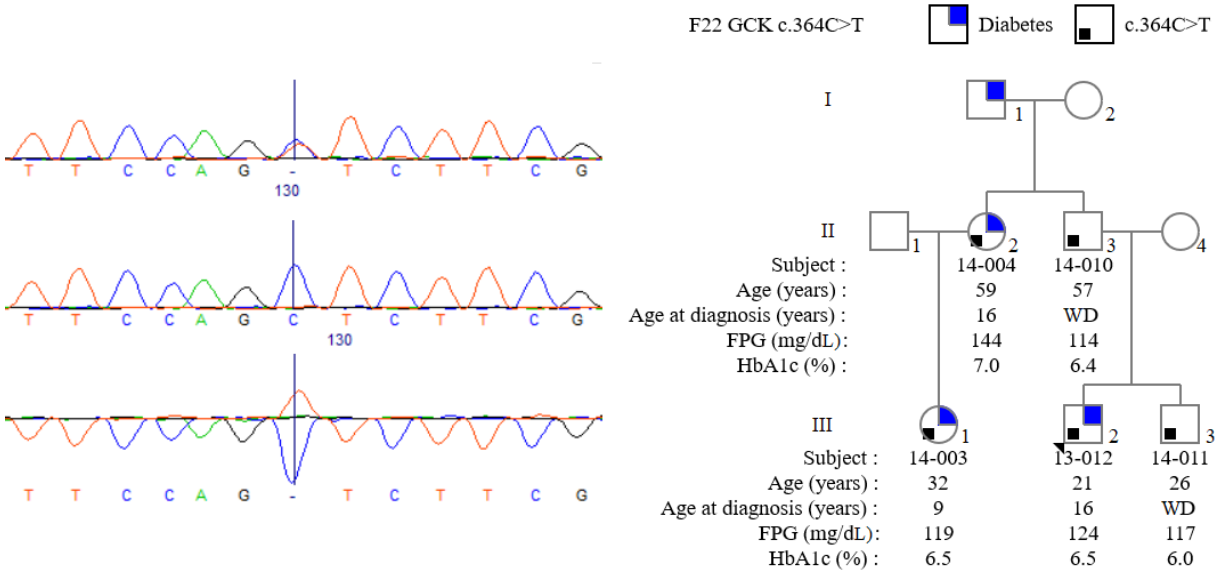


Figure 3.4: *GCK* c.364C>T proband chromatogram and pedigree. Proband chromatogram (top left), negative control (centre left), trace difference (bottom left) and pedigree (right). WD - without diagnosis. No data available for blank spaces.

The next variant, c.579+1_579+33del, can be classified as pathogenic (figure 3.5). This 33 bp deletion in intron 5 abolishes the donor splice site. Proband (III-1: 15-010) phenotype is suggestive of GCK-MODY, but cosegregation could not be ascertained as no relative samples were available. Though there is no population frequency data to our knowledge, this variant is associated with MODY in HGMD (CG942206) and has been previously reported as pathogenic (16,63).

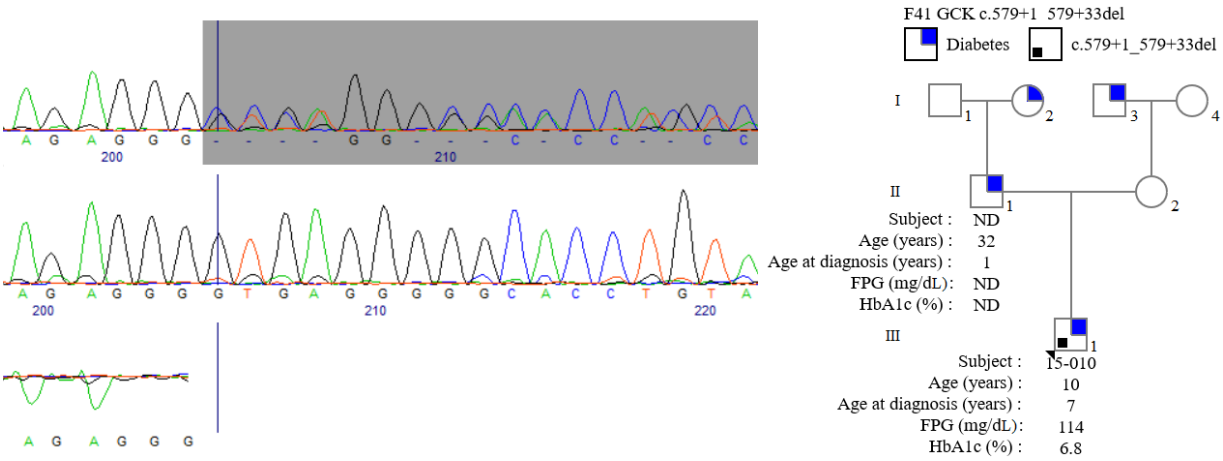


Figure 3.5: *GCK* c.579+1_579+33del proband chromatogram and pedigree. Proband chromatogram (top left), negative control (centre left), trace difference (bottom left) and pedigree (right). ND - no data available. No data available for blank spaces.

The third variant, c.766G>A (rs769268803) (figure 3.6), can also be classified as pathogenic. This exon 7 point substitution alters codon 256's first base (GAG>AAG), replacing the wild type aromatic glutamic acid with a basic lysine (p.Glu256Lys). Data on population frequency was only available on ExAC (MAF=0.00000824). This variant is classified as pathogenic on ClinVar and is associated with noninsulin-dependent diabetes on HGMD (CM930301). Additionally, this variant affects a conserved position and is predicted to be deleterious by MutationTaster (0.999999999995414), PolyPhen-2 (1.000; 0.986) and PROVEAN/SIFT (-3.73/0.000). Proband (III-1: 15-019) phenotype is consistent with GCK-MODY, but cosegregation could not be ascertained, as relative samples were unavailable. The c.766G>A variant has been previously reported has pathogenic (16,64–66).

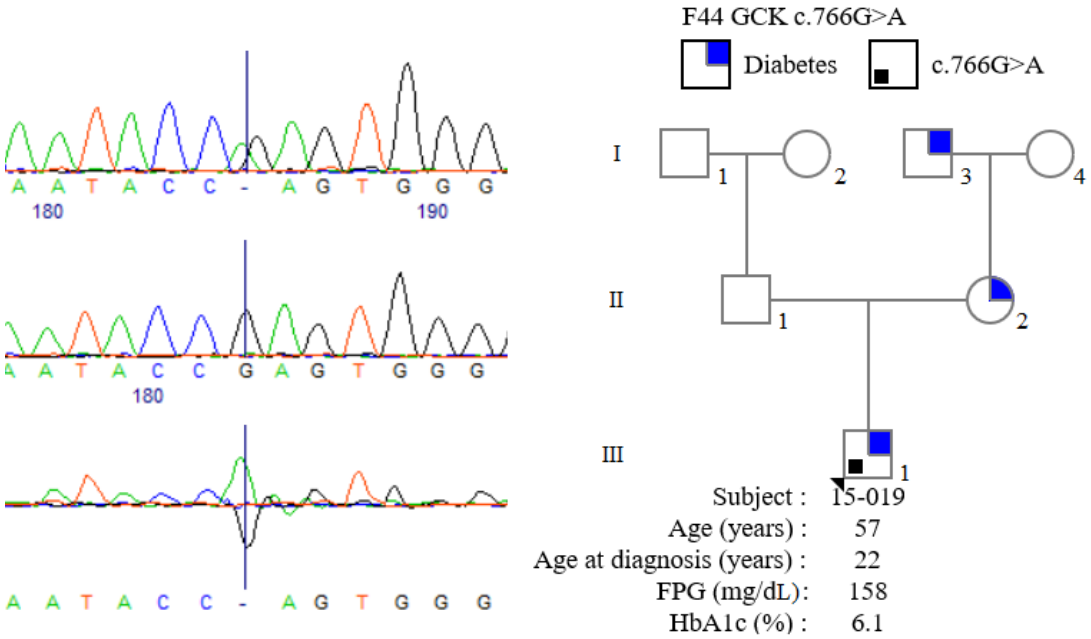


Figure 3.6: *GCK* c.766G>A proband chromatogram and pedigree. Proband chromatogram (top left), negative control (centre left), trace difference (bottom left) and pedigree (right). No data available for blank spaces.

The fourth and final variant, c.1268T>A (rs193922273) (figure 3.7), is also a missense variant (p.(Phe423Tyr)). This exon 10 point substitution alters codon 423's second nucleotide (TTC>TAC), changing the wild type aromatic phenylalanine to an also aromatic tyrosine. This variant can be

classified as likely pathogenic and is associated with MODY in HGMD (CM096954) and previous publications (16). The c.1268T>A variant affects a conserved position and is predicted to be deleterious by MutationTaster (0.999999544065252), PolyPhen-2 (0.998; 0.985) and SIFT (0.001). Proband (III-1: 15-011) and relative phenotypes are consistent with GCK-MODY, though c.1268T>A only cosegregates with diabetes in the proband's father. No population frequency data was available.

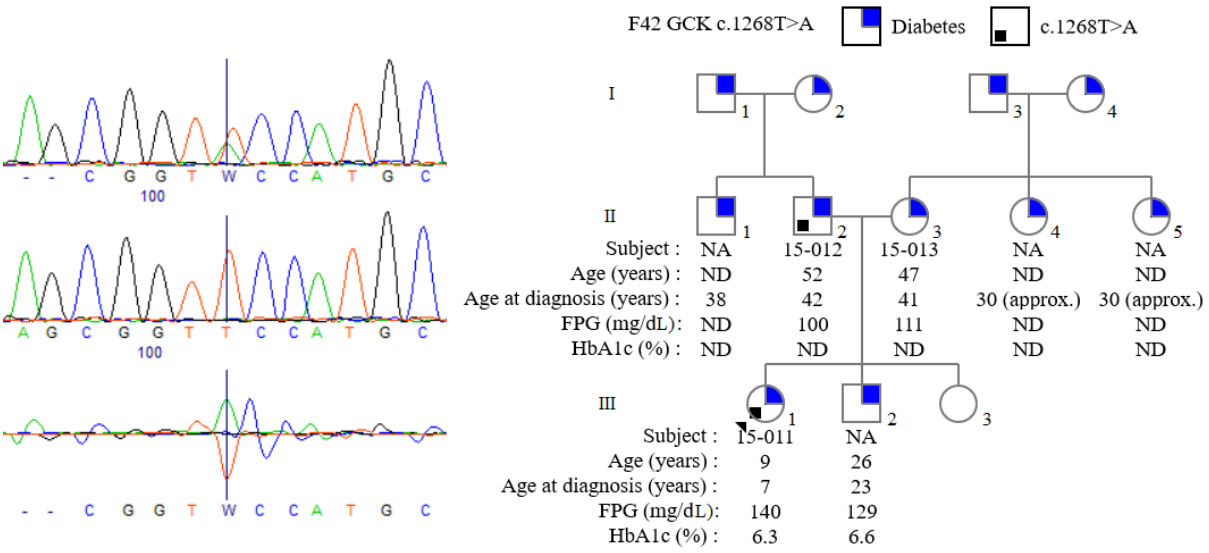


Figure 3.7: *GCK* c.1268T>A proband chromatogram and pedigree. Proband chromatogram (top left), negative control (15-013) (centre left), trace difference (bottom left) and pedigree (right). NA - not applicable, ND - no data available. No data available for blank spaces.

3.2.2 Sanger sequencing analysis: *HNF1A*

We detected 27 *HNF1A* sequence variants in 23 subjects, 18 probands and five relatives. Exon 1 and intron 2 were the most affected regions, with four variants present in each. No variants were detected in the putative promoter region, 5' UTR, and exons 2, 3, 5, 6 and 10 (figure 3.8).

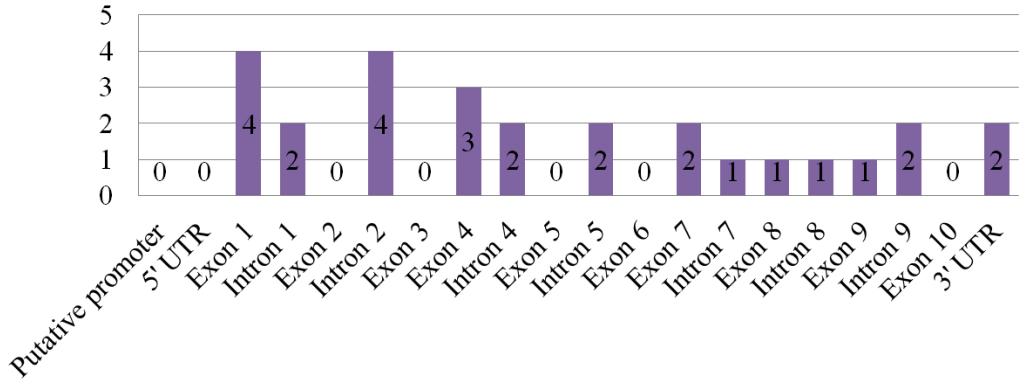


Figure 3.8: *HNF1A* sequence variants location. Distribution of *HNF1A* sequence variants throughout its genomic structure.

Again, most *HNF1A* variants were located in non-coding regions. However, coding region variants increased nearly three-fold, as 11 of 27 (40.74%) variants were located in coding regions. The two most common variants among probands, c.327-91G>A and c.1623+29T>C, were located in introns 1 and 8, and present in 15 and 13 probands, respectively. The third and fourth most common variants, c.864G>C and c.51C>G, were located in coding regions - exons 4 and 1 - and were present in 11 and 10 probands, respectively (figure 3.9). Furthermore, seven of the 11 (63.64%) coding region variants were located in exons 1 and 4, with the remaining four scattered through exons 7, 8 and 9 (figure 3.8).

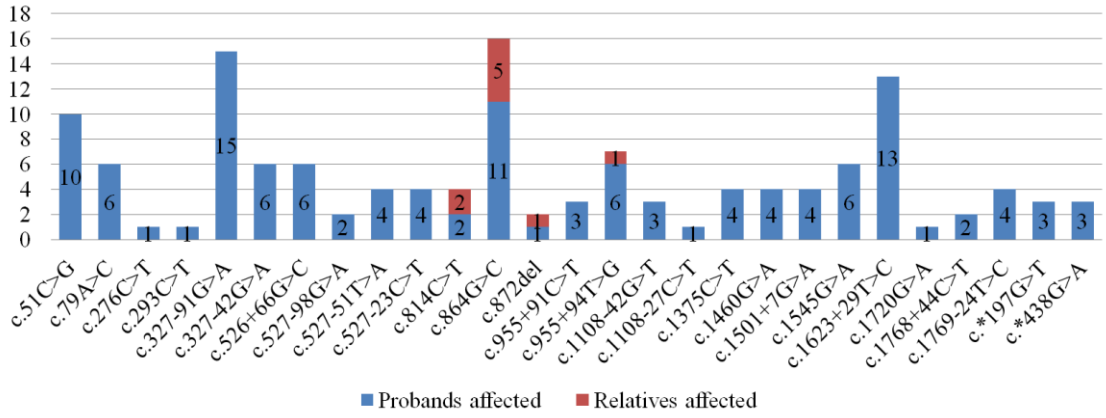


Figure 3.9: *HNF1A* sequence variants and affected subjects. Number of subjects affected by each *HNF1A* variant.

Two of 27 variants can be classified as pathogenic for HNF1A-MODY. The first, a missense variant, is located in exon 4. This point substitution in codon 272's first base (c.814C>T; CGC>TGC) (figure 3.10) ultimately replaces the wild type basic arginine with a nucleophilic cysteine (p.(Arg272Cys)). This variant can be classified as pathogenic as the nucleotide change takes place at a conserved position and the amino acid substitution is considered radical by MutationTaster. Also, c.814C>T is predicted to be deleterious by MutationTaster (0.999999999948752), PolyPhen-2 (1.000; 0.988) and PROVEAN/SIFT (-6.30/0.001). Proband (14-014 and 14-015) phenotypes are consistent with HNF1A-MODY, though cosegregation could only be ascertained for 14-015's family (F30). In this family, c.814C>T cosegregates with diabetes in 14-015's mother (15-016) and brother (15-018) but the father (15-017) does not have this variant, despite being diabetic (figure 3.11). No relative samples were available regarding proband 14-014. The c.814C>T variant is associated with T2DM in HGMD (CM991168) and MODY in literature (67). No population frequency data was available.

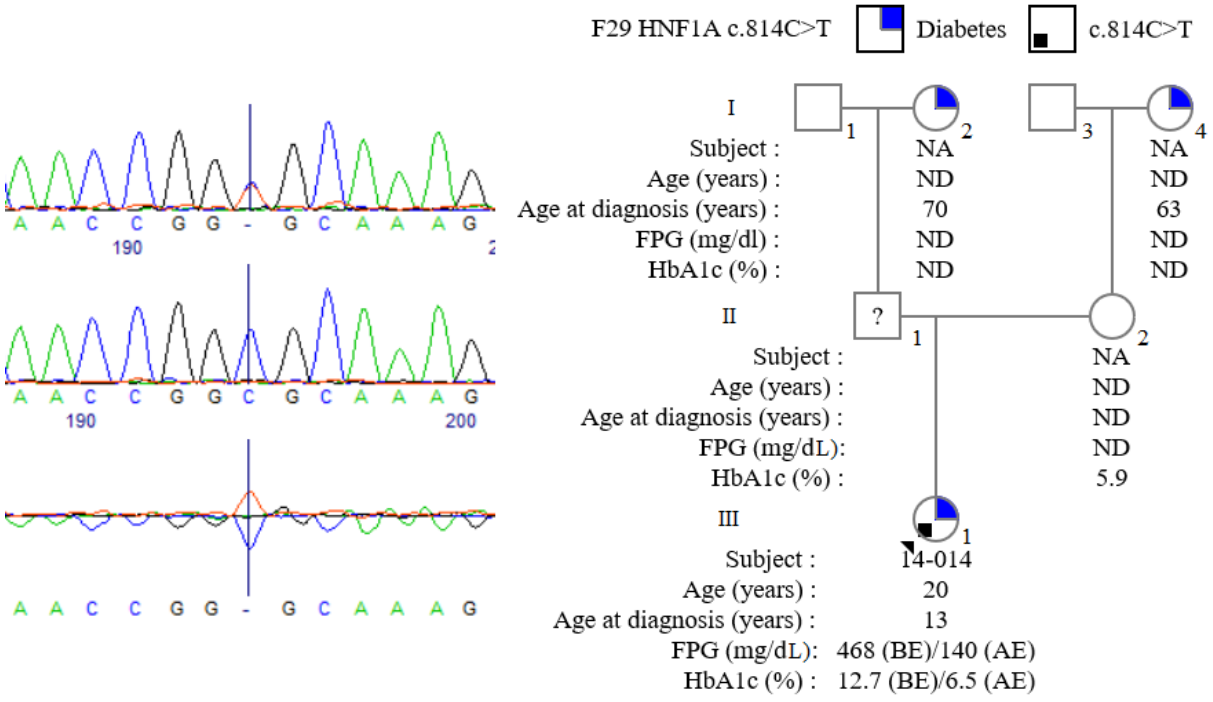


Figure 3.10: *HNF1A* c.814C>T proband (14-014) chromatogram and pedigree. Proband chromatogram (top left), negative control (centre left), trace difference (bottom left) and pedigree (right). Subject II-1 marked with "?" to represent he's unknown diabetes status, as he was going to draw blood for FPG evaluation. NA - not applicable, ND - no data available, BE - before enrollment, AE - at enrollment. No data available for blank spaces.

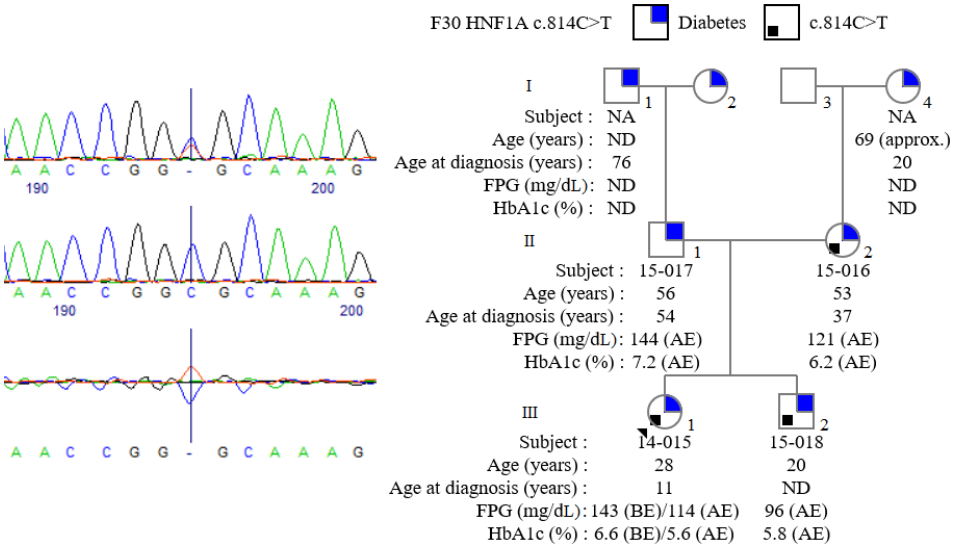


Figure 3.11: *HNF1A* c.814C>T proband (14-015) chromatogram and pedigree. Proband chromatogram (top left), negative control (15-017) (centre left), trace difference (bottom left) and pedigree (right). NA - not applicable, ND - no data available, BE - before enrollment, AE - at enrollment. No data available for blank spaces.

Also in exon 4, we detected a frameshift that can be classified as pathogenic. This variant, c.872del (figure 3.12), abolishes codon 291's second nucleotide (CCA), shifting the reading frame (p.(Pro291Glnfs*51)). Though the affected position does not seem to be conserved, MutationTaster predicts c.872del to be deleterious with a probability of 1. Proband (IV-1: 13-005) phenotype is consistent with MODY and c.872del cosegregates with diabetes in the mother (14-017). The father (14-018) does not have the variant nor diabetes (figure 3.12). This variant is associated with MODY in HGMD (CD972460) and literature (68) but there is no apparent ClinVar entry. No population frequency data was found.

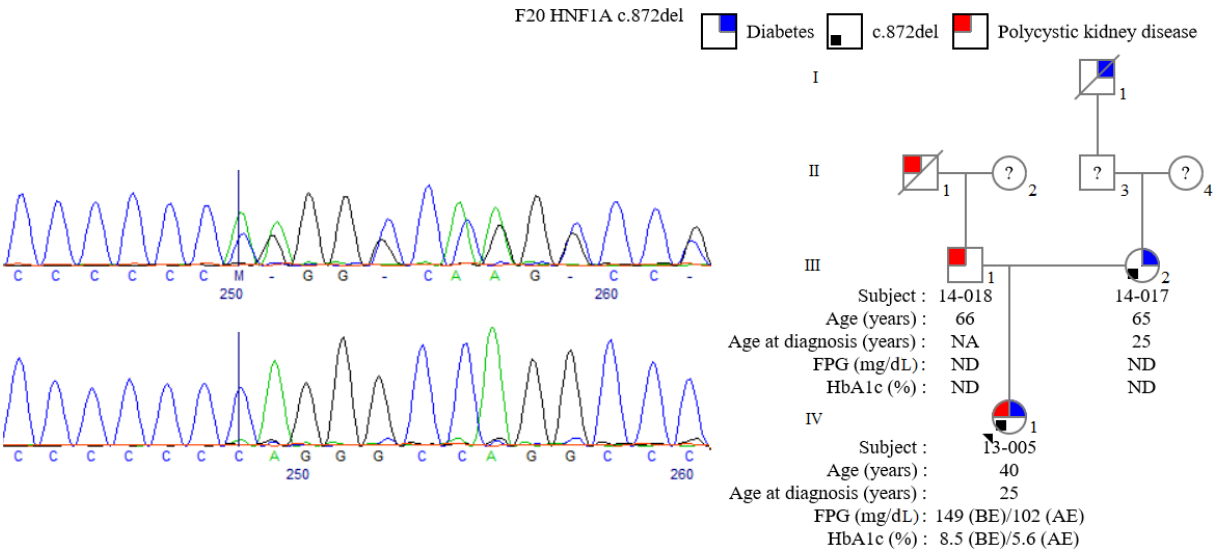


Figure 3.12: *HNF1A* c.872del proband chromatogram and pedigree. Proband chromatogram (top left), negative control (bottom left) and pedigree (right). Subjects marked with "?" have unknown diabetes status. NA - not applicable, ND - no data available, BE - before enrollment, AE - at enrollment. No data available for blank spaces.

In exon 9, a variant of uncertain significance was detected. This missense variant, c.1720G>A (rs1169305) (figure 3.13), affects codon 574's first nucleotide (GGC>AGC), replacing wild type small glycine with nucleophilic serine (p.(Gly574Ser)). Proband (III-2: 15-009) phenotype is consistent with diabetes but relative samples were not available for cosegregation study (figure 3.13). The c.1720G>A

variant is associated with T2DM in HGMD (CM020516). Proband is of African descent and data collected from the 1000 Genomes Browser points to a MAF of 0.05370000 in Africa and 0.0000 in Spain and Europe. According to Exome Variant Server, MAF in African-Americans is 0.038076.

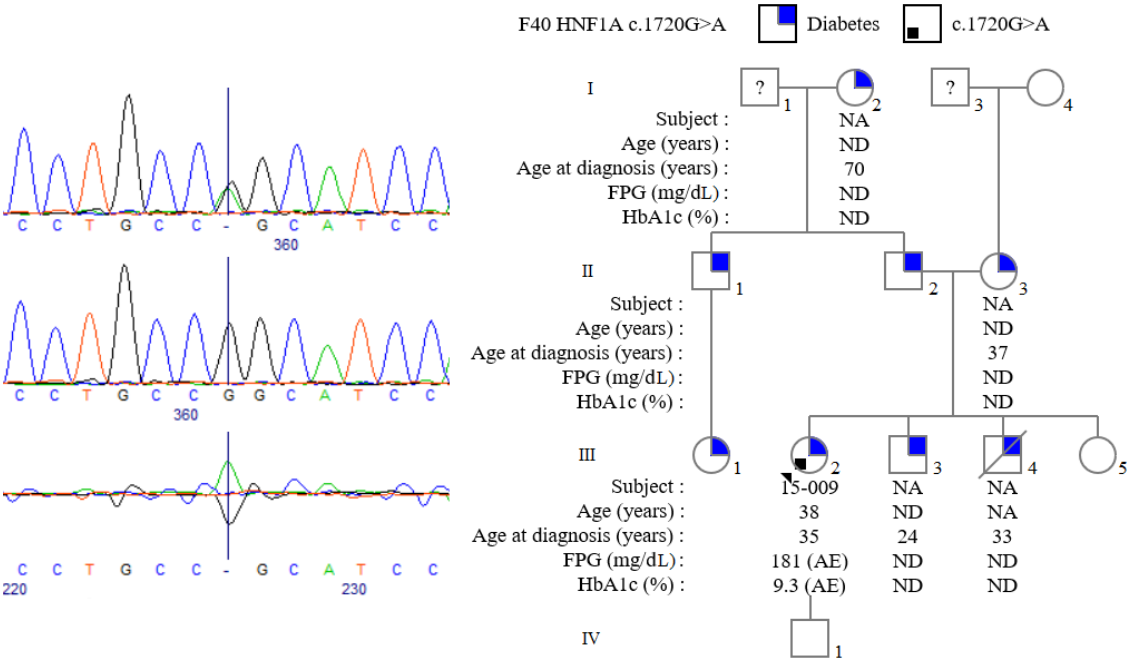


Figure 3.13: *HNF1A* c.1720G>A proband chromatogram and pedigree. Proband chromatogram (top left), negative control (centre left), trace difference (bottom left) and pedigree (right). Subjects marked with "?" have unknown diabetes status. NA - not applicable, ND - no data available, AE - at enrollment. No data available for blank spaces.

Three other relatively common missense variants were detected, but all three can be classified as benign or likely benign, in accordance with their ClinVar entries. Two are located in exon 1 (c.79A>C and c.293C>T) and the other is located in exon 7 (c.1460G>A). The benign c.79A>C (ATC>CTC) variant (rs1169288) (figure 3.14) replaces isoleucine with leucine (p.(Ile27Leu)), both hydrophobic amino acids. Though at a conserved position, the change itself is conservative. MutationTaster (0.0042728067953222), PolyPhen-2 (0.025; 0.022) and PROVEAN/SIFT (-0.92/0.082) all predict this variant to be benign. Among probands, five subjects are heterozygous and one subject is homozygous for c.79A>C, a variant associated with insulin resistance in HGMD (CM001349) but not with β-cell function (69). MAF is over 0.05 in Spain (0.3224), Europe (0.33899999) and Africa (0.0832000). The same is true for European-Americans (0.335156) and African-Americans (0.120744).

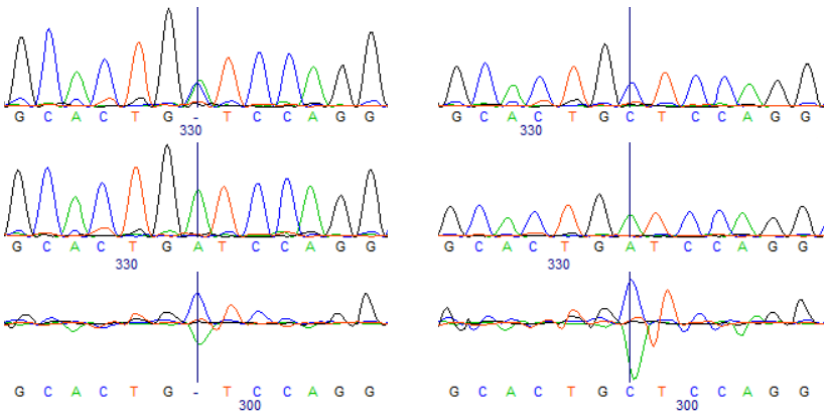


Figure 3.14: *HNF1A* c.79A>C proband chromatograms. Probands 15-008 (top left) and 15-009 (top right) chromatograms with respective negative controls (centre left and right) and trace differences (bottom left and right).

The likely benign c.293C>T (GCC>GTC) variant (rs1800574) (figure 3.15) replaces wild type small alanine with hydrophobic valine (p.(Ala98Val)), a moderately conservative change. Similarly, c.293C>T affects a conserved position, but MutationTaster (6.51464626319144e-06), PolyPhen-2 (0.328; 0.141) and PROVEAN/SIFT (-2.23/0.063) predict it to be benign. This variant is present in proband 15-019 and is associated with reduced serum C-peptide and insulin responses to an OGTT in HGMD (CM971442) and literature (70). MAF is under 0.05 in Spain (0.0047), Europe (0.0278000) and Africa (0.0030000). This remains true for European-Americans (0.025671) and African-Americans (0.004773).

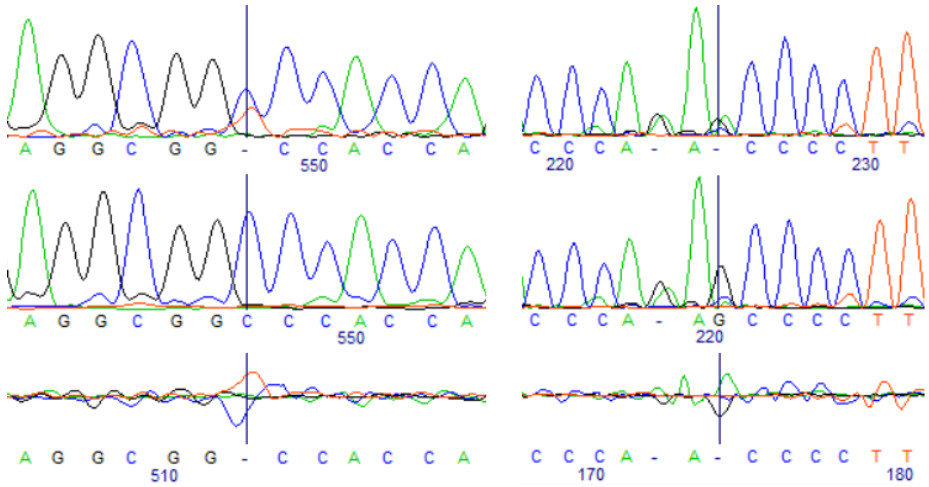


Figure 3.15: *HNF1A* c.293C>T and c.1460G>A proband chromatograms. Probands 15-019 (c.293C>T, top left) and 15-003 (c.1460G>A, top right), respective negative controls (centre left and right) and trace differences (bottom left and right).

The benign c.1460G>A (AGC>AAC) variant (rs2464196) (figure 3.15) replaces wild type nucleophilic serine for the amide asparagine (p.(Ser487Asn)), a conservative change. Again, c.1460G>A occurs at a conserved position, but MutationTaster (0.0158054424737475), PolyPhen-2 (0.014; 0.085) and PROVEAN/SIFT (-0.61/0.111) predict this variant to be benign. Among probands, four subjects are heterozygous for c.1460G>A. This variant is associated with increased risk for cardiovascular disease in HGMD (CM067474) and literature (71). MAF is over 0.05 in Spain (0.3178), Europe (0.3161) and Africa (0.112). The same holds true for European-Americans (0.311163) and African-Americans (0.122787).

3.2.3 MLPA analysis

A total of 17 proband samples had no detectable pathogenic/likely pathogenic sequence variants via Sanger sequencing and were therefore tested for CNV's via MLPA. Three samples yielded low quality results, meaning new samples for these probands will be necessary to perform new MLPA runs. No CNV's were detected in 11 of the tested samples. Two major deletions were identified in the remaining three proband samples, namely a heterozygous deletion of *GCK* exons 5 through 8 (c.484-?_1019+?del) in proband 15-014 (figure 3.16); and a heterozygous *HNF1B* deletion (c.1-?_1674+?del) in probands 13-006 and 14-013 (figure 3.17). Both deletions can be classified as pathogenic.

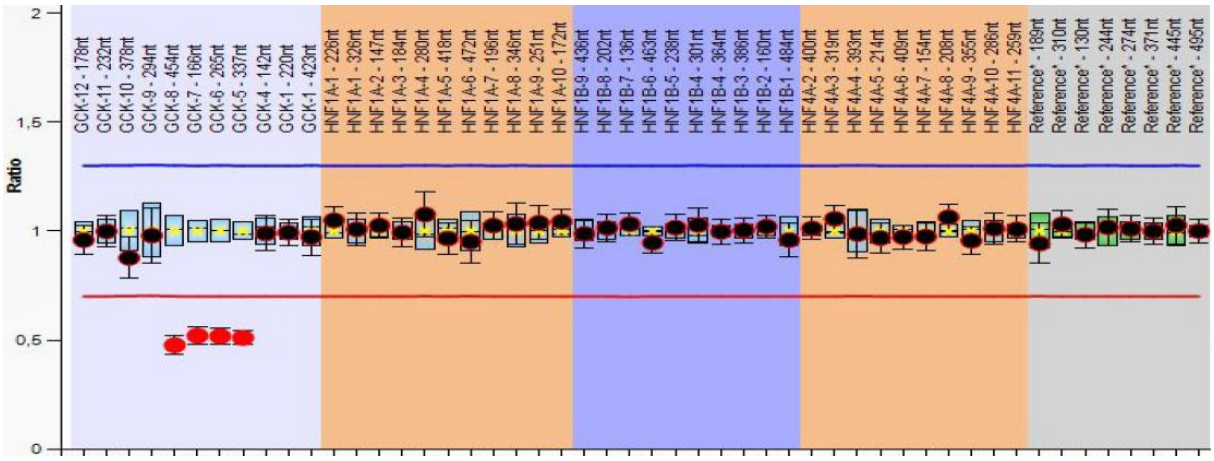


Figure 3.16: MLPA results for proband 15-014. Normalized fluorescence ratios are presented for all probes, including reference probes, with respective information regarding gene, exon and amplicon size. Area between blue and red lines corresponds to ratios between 0.7 and 1.3, indicating normal copy number. Ratios above 1.3 are indicative of duplication and ratios below 0.7 are indicative of deletion. Ratios between 1.3 and 1.65 are indicative of heterozygous duplication and ratios between 0.4 and 0.65 are indicative of heterozygous deletion.

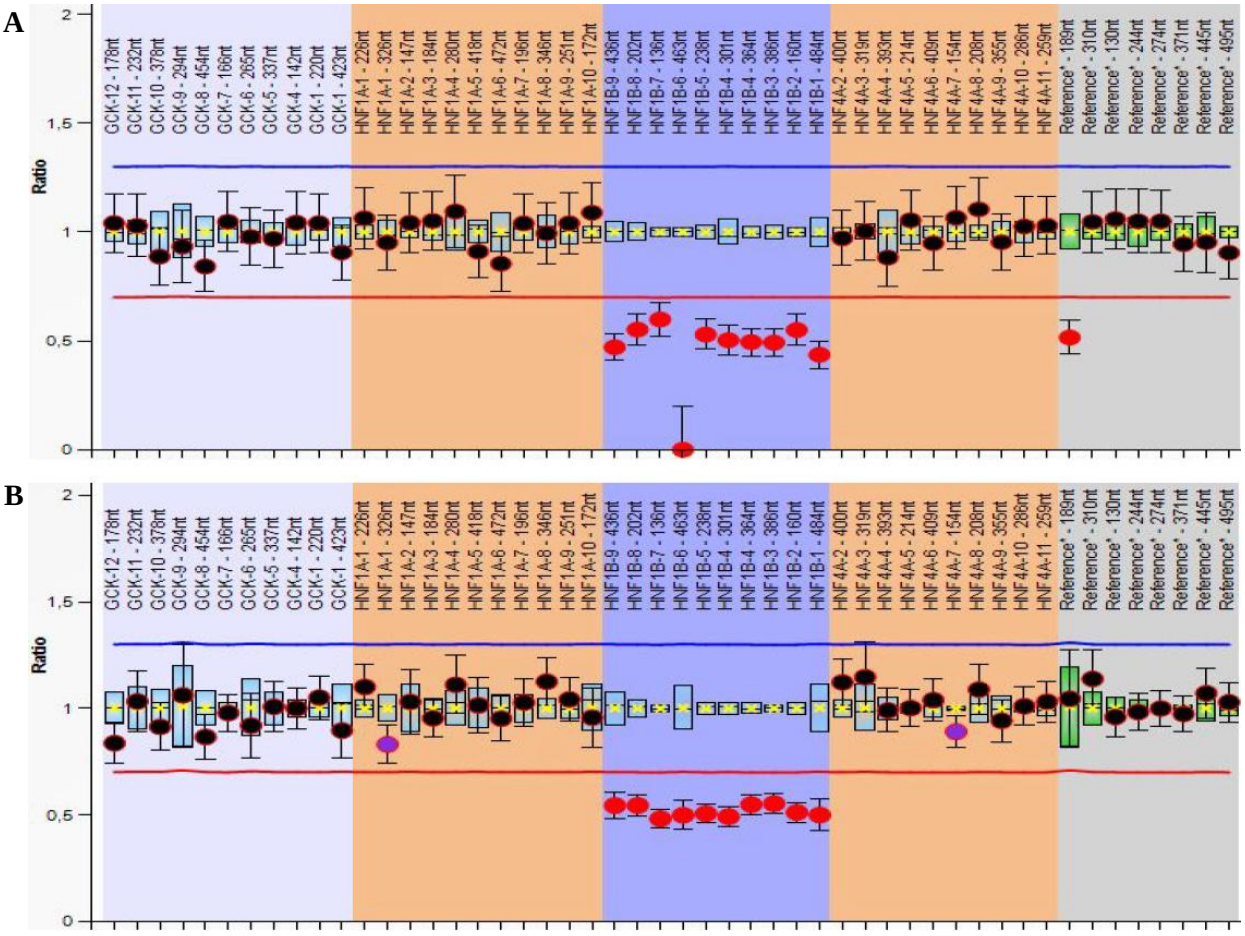


Figure 3.17: MLPA results for probands 13-006 (A) and 14-013 (B). Normalized fluorescence ratios are presented for all probes, including reference probes, with respective information regarding gene, exon and amplicon size. Area between blue and red lines corresponds to ratios between 0.7 and 1.3, indicating normal copy number. Ratios above 1.3 are indicative of duplication and ratios below 0.7 are indicative of deletion. Ratios between 1.3 and 1.65 are indicative of heterozygous duplication and ratios between 0.4 and 0.65 are indicative of heterozygous deletion. Ratios of zero are indicative of homozygous deletions.

4. Discussion

Overall, the study sample complies with MODY's diagnostic criteria. Probands are typically diagnosed before 25 years of age, even though three probands were diagnosed up to 10 years later. Relatives are generally diagnosed later, which may reflect the presence of atypical T2DM phenotypes, incomplete penetrance, environmental/behavioral factors delaying diabetes onset or lack of awareness regarding MODY among healthcare providers, particularly a few decades ago. Additionally, most probands presented with documented family history of diabetes, elevated plasma glucose (>100 mg/dL), no pancreatic autoantibodies and normal or low BMI, as some subjects are children or teenagers.

Interestingly, half the probands had normal HbA1c levels at enrollment and, though expected in HNF1A-MODY, vascular complications were not seen in these subjects, as the registered case of nephropathy was due to inherited polycystic kidney disease. This is most likely due to previous diagnosis and therapy implementation lowering circulating glucose and preventing future nefarious effects on the vascular system. In fact, most probands were on dietary therapy, insulin therapy, or both, and there is a clear reduction in both FPG and HbA1c from before study enrollment to the moment of enrollment, which occurs after healthcare intervention.

Unfortunately, data on every indicator was not available for every subject, meaning study sample characterization could be skewed and, as such, should be interpreted with caution. Likewise, considering subjects were referred from multiple hospitals, clinical assessment may have varied with the observing physician, as some T2DM phenotypes might have been mistaken for MODY.

In fact, some detected variants are associated with T2DM in HGMD and literature. The *GCK* c.1253+8C>T variant (rs2908274) in intron 9 was present in five probands - heterozygous in three, homozygous in two - and is associated with T2DM in HGMD (CS077759) and early onset T2DM in literature (72). However, this variant can be classified as benign for MODY and is considered polymorphic (*i.e.* benign) in literature (16). Also detected in *GCK*, the c.679+38T>C variant (rs2268574) in intron 6 is classified as benign (16), though it has been associated with GDM (73).

In *HNF1A*, T2DM association is more prevalent, which is expected, given *HNF1A*'s wider range of action. The previously mentioned c.79A>C, c.293C>T and c.1460G>A variants have been shown to decrease *HNF1A*'s transcriptional activity upon *GLUT2* (74) and have also been inconsistently associated with impaired glucose tolerance and/or an increased risk of T2DM (75). The c.79A>C variant is located in the dimerization domain, a highly conserved region of great functional significance (74), but does not seem to be associated with altered C-peptide and insulin release (70). However, it has been associated with insulin resistance, albeit via an unknown mechanism. Considering IGT is an insulin resistance state, the c.79A>C variant might be a risk factor or play a role in T2DM pathogenesis (69,76). It is also worth mentioning this variant is not associated with β-cell function (69). The c.79A>C variant is also associated with increased risk of subclinical coronary atherosclerosis and coronary heart disease, and with elevated LDL (71).

The c.293C>T variant is much less frequent than the other two and, despite its likely benign classification, seems to be associated with reduced serum C-peptide and insulin responses to an oral glucose challenge (70). The c.79A>C and c.1460G>A variants have similar MAFs and have been shown to be in linkage disequilibrium (74). This should explain why the latter was also found to not be associated with altered C-peptide and insulin release (70). In addition to the T2DM link (77), and like the c.79A>C variant, the c.1460G>A variant has also been associated with increased cardiovascular disease risk and high LDL (71). The up- and downstream c.1375C>T (p.(=); rs2259820) and c.1501+7G>A (rs2464195) variants, in exon and intron 7, respectively, were also detected and show

very similar MAFs. The c.1375C>T variant is also associated with T2DM (77) and it would not be surprising if the intronic c.1501+7G>A variant shared this feature, as these three are likely in linkage disequilibrium. Furthermore, these three are classified as benign/likely benign in ClinVar.

In intron 2, the relatively common c.527-51T>A (rs2071190) variant was detected in four subjects (one homozygous, three heterozygous) and is significantly associated with T2DM in both literature (78) and HGMD (CS066655). In exon 4, the also common c.864G>C (p.(=); rs56348580) variant was detected in a total of 16 subjects (eight homozygous, eight heterozygous) and, though classified as benign for MODY, it has been associated with T2DM in HGMD (CM067044) and literature (79).

Most variants were located in non-coding regions. *GCK* variants seem scattered throughout the gene, as previously reported (16), with the exception of intron 8, which presented the greatest allelic diversity. In *HNF1A*, more than half of variants are located between exon 1 and intron 4, in agreement with previous reports (15,31). *HNF1A* variants comprised the majority of detected sequence variants, reflecting its great allelic heterogeneity (31). However, pathogenic or likely pathogenic variants for MODY were proportionately more common in *GCK*.

Regarding these variants, in *GCK*, no additional information on c.364C>T was found. Absence of frequency data hints at a very low MAF, though this variant has been previously reported in two New Zealand families (16) and one Japanese family (80). The c.579+1_579+33del variant has been previously reported in three families, two French (16,63) and one Brazilian (81), and should equally have very low MAF, as there is seemingly no frequency data. This deletion likely results in a null variant, since both intron 5 donor splice site consensus bases (GT) are lost, but further RNA studies are necessary to understand the molecular consequences. The c.766G>A (p.Glu256Lys) variant has been previously reported in 11 families across Sweden, Spain and France (16,64,65), and seemingly decreases enzymatic activity. Glu256 is located in GCK's active site and, in its open form, is implicated in glucose binding by forming hydrogen-bond interactions with hydroxyls of bound glucose. Mutation to Lysine decreases maximal velocity (V_{max}) with no significant change in K_m for glucose, and should reduce glucose affinity by removing the hydrogen-bond (65). The Glu256Lys variant induces conformational changes in GCK's active site as well as the whole structure, resulting in decreased glucose binding and downstream loss of catalytic activity, thus explaining the hyperglicemic phenotype (66). Finally, the c.1268T>A variant may also have very low MAF, as no frequency data was found, though it was previously reported in multiple families (16), some in Portugal (82) and Brazil (83).

In *HNF1A*, the c.814C>T variant has been previously reported in Japan (67), Italy (84) and Iceland (85). It is located in the DNA-binding domain and renders the protein unable to bind DNA, leading to a mutant with no transactivating activity. The resulting p.(Arg272Cys) protein acts in a dominant-negative fashion, as HNF1A operates as a homodimer. The same study also reported a reduced insulin and glucagon secretory responses to arginine (67), highlighting α- and β-cell dysfunction.

The c.872del variant is located in a transactivation domain proline rich region. This variant is predicted to lead to the synthesis of a truncated protein, due to an alternative termination codon arising 51 codons downstream of the frameshift (p.(Pro291Glnfs*51)) (68). As the dimerization domain remains unchanged, it is possible this variant still forms a homodimer with the wild type protein, or with HNF1B, meaning it would act in a dominant-negative manner (85). It should be noted that frameshifts have been previously reported in the same position, one nucleotide downstream or one codon upstream, due to insertions (64), duplications (15,86) and deletions (85), respectively. This along with the fact that the wild type exonic region where these variants occur has eight C nucleotides

in tandem (figure 3.12) supports the hypothesis that this is a mutation hotspot resulting from replication slippage (85). Furthermore, subjects affected by the c.872del variant are homozygous for the benign but T2DM-associated c.864G>C, which adds another C nucleotide to this region.

Finally, the c.1720G>A variant has been classified as both pathogenic and benign in ClinVar. Despite being very rare in Europe and associated with T2DM and atypical diabetes in African-American children (87), this variant occurs at a poorly conserved position, has a MAF over 5% in Africa and is predicted to be benign by PolyPhen-2 (0.002; 0.003) and PROVEAN/SIFT (-0.08/0.779). Given the subjects ancestry and conflicting data, it is not possible to conclude on this variant's pathogenicity.

Of the six subjects (three probands and three relatives) with pathogenic *HNF1A* variants, four were on insulin therapy and one on OHA, underlining their poor glycemic control. As most of them were on therapy at enrollment, mean FPG was just mildly elevated (114.6 ± 17.26 mg/dL) and HbA1c was normal (5.94 ± 0.40%). Age at diagnosis was also lower for exon 4 pathogenic variant carriers versus exon 9 variant carriers, as expected (3), though the number of subjects is insufficient to reach solid conclusions. Age at diagnosis was unexpectedly higher in subjects with frameshift variants in comparison with missense variants (25 versus 20.33 ± 14.47 years) though, again, the number of subjects was insufficient to reach solid conclusions.

Of the ten subjects (five probands and five relatives) with pathogenic/likely pathogenic *GCK* variants, three were on diet, three other on OHA and the remaining four were not on treatment or had no available data. With the correct diagnosis, and considering GCK-MODY individuals usually retain good homeostatic control over blood glucose, it might benefit these subjects to follow a dietary therapy and pursue an active lifestyle. These subjects were diagnosed at different stages of their lives, either in their youth (age 7 and 9 years), their teens (16 years), young adulthood (at 22 years) and adulthood (at 42 years), as expected in GCK-MODY (4).

Regarding the MLPA results, the heterozygous *GCK* exons 5 through 8 deletion should result in a null variant, decreasing the phosphorylation rate which, duly compensated by the wild type allele, would generate the characteristic GCK-MODY phenotype (4,16), even though the breaking points are not known. There was little clinical information available on proband 15-014, who was diagnosed at 16 years of age and presented with family history of diabetes though, at the time, relative samples were unavailable for cosegregation study. At enrollment, FPG was mildly elevated (126 mg/dL) and HbA1c was 5.9%, in line with typical GCK-MODY phenotype. There was no data on patient treatment, although it is likely this proband was put on a diet. Partial *GCK* deletions seem to be a rare cause of GCK-MODY (16) and there is seemingly no information concerning a deletion of exons 5 through 8 - and its effects at the protein level - in ClinVar, dbVar, HGMD and PubMed, meaning we could be in the presence of a novel pathogenic mutation. Also, considering the *GCK* c.766G>A (p.Glu256Lys) variant in exon 7 affects the active site, decreasing glucose binding and catalytic activity, it stands to reason this multiexon deletion would have at least similar consequences (24,66).

Conversely, *HNF1B* deletions are frequent (4,11,88,89). There are several entries concerning MODY-associated whole gene deletions among HGMD, ClinVar, dbVar and PubMed, with many of these deletions spanning other genes at 17q12. Heterozygous *HNF1B* deletion was detected in two probands but, as breaking points are unknown, it is possible more genes are affected by these deletions. In fact, *HNF1B* deletion has been reported as the same genetic disorder as 17q12 deletion (90). Proband 13-006 has no apparent family history of diabetes. However, this could be a *de novo* mutation, as they are frequent (4). Unfortunately, relative samples were unavailable and cosegregation study was not done. In opposition, proband 14-013 has two diabetic grandparents, one maternal and the other paternal but,

again, relative samples were not available. Proband 13-006 has a history of extremely high FPG and HbA1c - 600 mg/dL and 14.1% around the time of diagnosis, but 78 mg/dL and 5.5% at enrollment - while 14-013 shows just mild hyperglicemia (120 mg/dL) and normal HbA1c (5.8%), before and at enrollment. Insulin resistance is typical of HNF1B-MODY and patients are often on insulin treatment at early stages (91). In this case, both probands were on diet and insulin therapy. Additionally, though it is not known if the deleted region is exactly the same in both probands, HNF1B-MODY phenotypes are notorious for their heterogeneity (91), even in cases with the same mutation (44,48). Amazingly, both probands had no record of any renal abnormalities or other abnormalities of any kind. Renal cysts are a typical HNF1B-MODY feature (44), though their absence has been reported (91,92). However, it is possible these and other abnormalities were not thoroughly screened for.

These whole gene deletions are pathogenic for HNF1B-MODY, as evidenced by several HGMD entries (CG055897, CG094325 and CG106852) (88,89), and seem to result in disease via haploinsufficiency (93,94). Considering HNF1B typically acts as a homodimer, hemizygous individuals will most likely be unable to produce enough homodimers to assure wild type phenotype expression.

In the particular case of proband 13-006, normalized fluorescence ratios for two probes call for further analysis. The 189 nucleotide reference probe ratio is below normal, seemingly indicating a heterozygous deletion of the corresponding region. However, copy number changes of reference probes are unlikely and this particular probe is located within, or close to, a very strong CpG island. This low ratio could result from CpG islands' high GC content, which would hinder sample denaturation and subsequent probe hybridisation, ligation and amplification, generating a false positive result. These denaturation problems would also explain the apparent homozygous deletion suggested by HNF1B-6 probe's ratio (figure 3.17 - A). Actually, poor sample denaturation can result in false positives even when ratios suggest the apparent deletion of adjacent genomic regions (62), such as the *GCK* exons 5 through 8 deletion. Another possibility would be the high GC content in the sequence adjacent to HNF1B-6 probe's ligation site. As for the three probands whose samples generated low quality results, new samples are need to retest. These three samples had low DNA concentration and/or fragmented DNA, which could explain the obtained results.

Additionally, it should be noted the MLPA technique has several limitations. Mutations or polymorphisms present in sequences detected by MLPA probes can reduce relative fluorescence peak heights, even when not located at the exact ligation site between each probe's oligonucleotide pair, therefore generating false positives. Furthermore, MLPA is a sensitive technique and its performance relies on operator proficiency. As specified in the MLPA protocol and probemix description (62), changes in sample purity and quantities or experimental conditions can also affect peak height for some probes. As such, MLPA results need confirmation by other methods. Considering the reported deletions' breaking points are unknown, long range PCR would serve both purposes, confirming the reported results and clarifying the breaking points locations.

In conclusion, 10 out of 24 probands were identified as carriers of pathogenic/likely pathogenic MODY variants: five with GCK-MODY, three with HNF1A-MODY and two with HNF1B-MODY. Of the remaining 14 probands without any detected pathogenic/likely pathogenic variants, 11 had at least one T2DM associated variant and six of the 11 had T2DM associated variants on *GCK* and *HNF1A*. Thus, this study highlights the need for comprehensive and documented patient evaluation, including genetic testing, as a correct diagnosis impacts patient treatment and quality of life. Future work should focus on improving the current *status quo*, further educating healthcare providers on MODY's characteristics and diagnosis, as well as the benefits of genetic screening in patient

management. Development, improvement and implementation of online tools, such as diagnostic algorithms (17), can also aid healthcare providers. Furthermore, it is necessary to reach more individuals within national territory, screen more MODY genes, and establish a control panel to improve variant ascertainment and create a representative report of MODY's prevalence and features in the Portuguese population. Functional studies to determine the biological consequences of different variants, at mRNA and protein levels, are essential to better understand the underlying processes of MODY's subtypes. The creation of dedicated units to screen and monitor target populations in defined geographical areas (north, centre, south, Azores and Madeira archipelagos) would help in establishing a national database with anamnesis, molecular genetic and biochemical information that would ultimately improve patient diagnosis and management.

5. References

1. Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 2009 Jan;32(Suppl 1):S62–7.

2. Tattersall RB. Mild Familial Diabetes with Dominant Inheritance. QJM. 1974 Apr 1;43(2):339–57.

3. Ellard S, Bellanné-Chantelot C, Hattersley AT. Best practice guidelines for the molecular genetic diagnosis of maturity-onset diabetes of the young. Diabetologia. 2008 Apr;51(4):546–53.

4. McDonald TJ, Ellard S. Maturity onset diabetes of the young: identification and diagnosis. Ann Clin Biochem Int J Biochem Lab Med. 2013 Sep 1;50(5):403–15.

5. Anik A, Catli G, Abaci A, Bober E. Maturity-onset diabetes of the young (MODY): an update. J Pediatr Endocrinol Metab. 2015 Mar;28(3–4):251–63.

6. Kim SH. Maturity-Onset Diabetes of the Young: What Do Clinicians Need to Know? Diabetes Metab J. 2015 Dec;39(6):468–77.

7. Pihoker C, Gilliam LK, Ellard S, Dabelea D, Davis C, Dolan LM, et al. Prevalence, Characteristics and Clinical Diagnosis of Maturity Onset Diabetes of the Young Due to Mutations in HNF1A, HNF4A, and Glucokinase: Results From the SEARCH for Diabetes in Youth. J Clin Endocrinol Metab. 2013 Jun 14;98(10):4055–62.

8. Thanabalasingham G, Pal A, Selwood MP, Dudley C, Fisher K, Bingley PJ, et al. Systematic Assessment of Etiology in Adults With a Clinical Diagnosis of Young-Onset Type 2 Diabetes Is a Successful Strategy for Identifying Maturity-Onset Diabetes of the Young. Diabetes Care. 2012 Jun;35(6):1206–12.

9. Kavvoura FK, Owen KR. Maturity onset diabetes of the young: clinical characteristics, diagnosis and management. Pediatr Endocrinol Rev PER. 2012 Jan;10(2):234–42.

10. Schober E, Rami B, Grabert M, Thon A, Kapellen T, Reinehr T, et al. Phenotypical aspects of maturity-onset diabetes of the young (MODY diabetes) in comparison with Type 2 diabetes mellitus (T2DM) in children and adolescents: experience from a large multicentre database. Diabet Med. 2009 May 1;26(5):466–73.

11. Yorifuji T, Fujimaru R, Hosokawa Y, Tamagawa N, Shiozaki M, Aizu K, et al. Comprehensive molecular analysis of Japanese patients with pediatric-onset MODY-type diabetes mellitus. Pediatr Diabetes. 2012 Feb 1;13(1):26–32.

12. Shields BM, Hicks S, Shepherd MH, Colclough K, Hattersley AT, Ellard S. Maturity-onset diabetes of the young (MODY): how many cases are we missing? Diabetologia. 2010 May 25;53(12):2504–8.

13. Xu JY, Dan QH, Chan V, Wat NMS, Tam S, Tiu SC, et al. Genetic and clinical characteristics of maturity-onset diabetes of the young in Chinese patients. Eur J Hum Genet. 2004 Dec 15;13(4):422–7.

14. Hwang JS, Shin CH, Yang SW, Jung S young, Huh N. Genetic and clinical characteristics of Korean maturity-onset diabetes of the young (MODY) patients. Diabetes Res Clin Pract. 2006 Oct;74(1):75–81.

15. Colclough K, Bellanne-Chantelot C, Saint-Martin C, Flanagan SE, Ellard S. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha and 4 alpha in maturity-onset diabetes of the young and hyperinsulinemic hypoglycemia. Hum Mutat. 2013 May;34(5):669–85.

16. Osbak KK, Colclough K, Saint-Martin C, Beer NL, Bellanné-Chantelot C, Ellard S, et al. Update on mutations in glucokinase (GCK), which cause maturity-onset diabetes of the young, permanent neonatal diabetes, and hyperinsulinemic hypoglycemia. Hum Mutat. 2009 Nov 1;30(11):1512–26.

17. Shields BM, McDonald TJ, Ellard S, Campbell MJ, Hyde C, Hattersley AT. The development and validation of a clinical prediction model to determine the probability of MODY in patients with young-onset diabetes. Diabetologia. 2012 May;55(5):1265–72.

18. McDonald TJ, Colclough K, Brown R, Shields B, Shepherd M, Bingley P, et al. Islet autoantibodies can discriminate maturity-onset diabetes of the young (MODY) from Type 1 diabetes. Diabet Med J Br Diabet Assoc. 2011 Sep;28(9):1028–33.

19. Pearson ER, Starkey BJ, Powell RJ, Gribble FM, Clark PM, Hattersley AT. Genetic cause of hyperglycaemia and response to treatment in diabetes. The Lancet. 2003 Oct 18;362(9392):1275–81.

20. Parker HE, Habib AM, Rogers GJ, Gribble FM, Reimann F. Nutrient-dependent secretion of glucose-dependent insulinotropic polypeptide from primary murine K cells. Diabetologia. 2009 Feb;52(2):289–98.

21. Reimann F, Habib AM, Tolhurst G, Parker HE, Rogers GJ, Gribble FM. Glucose Sensing in L Cells: A Primary Cell Study. Cell Metab. 2008 Dec 6;8(6–3):532–9.

22. Iynedjian PB. Molecular Physiology of Mammalian Glucokinase. Cell Mol Life Sci. 2009 Jan 1;66(1):27.

23. Fu Z, Gilbert ER, Liu D. Regulation of Insulin Synthesis and Secretion and Pancreatic Beta-Cell Dysfunction in Diabetes. Curr Diabetes Rev. 2013 Jan 1;9(1):25–53.

24. Kamata K, Mitsuya M, Nishimura T, Eiki J, Nagata Y. Structural Basis for Allosteric Regulation of the Monomeric Allosteric Enzyme Human Glucokinase. Structure. 2004 Mar 1;12(3):429–38.

25. Baldini SF, Steenackers A, Olivier-Van Stichelen S, Mir A-M, Mortuaire M, Lefebvre T, et al. Glucokinase expression is regulated by glucose through O-GlcNAc glycosylation. Biochem Biophys Res Commun. 2016 Sep 16;478(2):942–8.

26. Velho G, Petersen KF, Perseghin G, Hwang JH, Rothman DL, Pueyo ME, et al. Impaired hepatic glycogen synthesis in glucokinase-deficient (MODY-2) subjects. J Clin Invest. 1996 Oct 15;98(8):1755–61.

27. Byrne MM, Sturis J, Clément K, Vionnet N, Pueyo ME, Stoffel M, et al. Insulin secretory abnormalities in subjects with hyperglycemia due to glucokinase mutations. J Clin Invest. 1994 Mar;93(3):1120–30.

28. Liu S, Ammirati MJ, Song X, Knafels JD, Zhang J, Greasley SE, et al. Insights into Mechanism of Glucokinase Activation. J Biol Chem. 2012 Apr 20;287(17):13598–610.

29. Stride A, Shields B, Gill-Carey O, Chakera AJ, Colclough K, Ellard S, et al. Cross-sectional and longitudinal studies suggest pharmacological treatment used in patients with glucokinase mutations does not alter glycaemia. Diabetologia. 2014;57(1):54–6.

30. Spyer G, Hattersley AT, Sykes JE, Sturley RH, MacLeod KM. Influence of maternal and fetal glucokinase mutations in gestational diabetes. Am J Obstet Gynecol. 2001 Jul 1;185(1):240–1.

31. Bellanne-Chantelot C, Carette C, Riveline J-P, Valero R, Gautier J-F, Larger E, et al. The type and the position of HNF1A mutation modulate age at diagnosis of diabetes in patients with maturity-onset diabetes of the young (MODY)-3. Diabetes. 2008 Feb;57(2):503–8.

32. Harries LW, Brown JE, Gloyn AL. Species-Specific Differences in the Expression of the HNF1A, HNF1B and HNF4A Genes. PLOS ONE. 2009 Nov 16;4(11):e7855.

33. Harries LW, Ellard S, Stride A, Morgan NG, Hattersley AT. 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. Hum Mol Genet. 2006 Jul 15;15(14):2216–24.

34. Odom DT, Zizlsperger N, Gordon DB, Bell GW, Rinaldi NJ, Murray HL, et al. Control of Pancreas and Liver Gene Expression by HNF Transcription Factors. Science. 2004 Feb 27;303(5662):1378–81.

35. Servitja J-M, Pignatelli M, Maestro MÁ, Cardalda C, Boj SF, Lozano J, et al. Hnf1α (MODY3) Controls Tissue-Specific Transcriptional Programs and Exerts Opposed Effects on Cell Growth in Pancreatic Islets and Liver. Mol Cell Biol. 2009 Jun 1;29(11):2945–59.

36. Stride A, Vaxillaire M, Tuomi T, Barbetti F, Njølstad PR, Hansen T, et al. The genetic abnormality in the beta cell determines the response to an oral glucose load. Diabetologia. 2002 Mar 1;45(3):427–35.

37. Stride A, Ellard S, Clark P, Shakespeare L, Salzmann M, Shepherd M, et al. β-Cell Dysfunction, Insulin Sensitivity, and Glycosuria Precede Diabetes in Hepatocyte Nuclear Factor-1α Mutation Carriers. Diabetes Care. 2005 Jul 1;28(7):1751–6.

38. Thomas H, Jaschkowitz K, Bulman M, Frayling TM, Mitchell SM, Roosen S, et al. A distant upstream promoter of the HNF-4alpha gene connects the transcription factors involved in maturity-onset diabetes of the young. Hum Mol Genet. 2001 Sep 15;10(19):2089–97.

39. Harries LW, Locke JM, Shields B, Hanley NA, Hanley KP, Steele A, et al. The Diabetic Phenotype in HNF4A Mutation Carriers Is Moderated By the Expression of HNF4A Isoforms From the P1 Promoter During Fetal Development. Diabetes. 2008 Jun 1;57(6):1745–52.

40. Eeckhoute J, Moerman E, Bouckenooghe T, Lukoviak B, Pattou F, Formstecher P, et al. Hepatocyte Nuclear Factor 4α Isoforms Originated from the P1 Promoter Are Expressed in Human Pancreatic β-Cells and Exhibit Stronger Transcriptional Potentials than P2 Promoter-Driven Isoforms. Endocrinology. 2003 May 1;144(5):1686–94.

41. Chandra V, Huang P, Potluri N, Wu D, Kim Y, Rastinejad F. Multi-Domain Integration in the Structure of the HNF4α Nuclear Receptor Complex. Nature. 2013 Mar 21;495(7441):394–8.

42. Parviz F, Matullo C, Garrison WD, Savatski L, Adamson JW, Ning G, et al. Hepatocyte nuclear factor 4alpha controls the development of a hepatic epithelium and liver morphogenesis. Nat Genet. 2003 Jul;34(3):292–6.

43. Ribeiro A, Pastier D, Kardassis D, Chambaz J, Cardot P. Cooperative Binding of Upstream Stimulatory Factor and Hepatic Nuclear Factor 4 Drives the Transcription of the Human Apolipoprotein A-II Gene. J Biol Chem. 1999 Jan 15;274(3):1216–25.

44. Edghill EL, Bingham C, Ellard S, Hattersley AT. Mutations in hepatocyte nuclear factor-1β and their related phenotypes. J Med Genet. 2006 Jan;43(1):84–90.

45. Aboudehen K, Patel V, Kim M, Xie Y, McDonald J, Gordillo R, et al. Identification of hepatocyte nuclear factor-1β gene regulatory networks in mouse kidney cells (690.9). FASEB J. 2014 Apr 1;28(1 Supplement):690.9.

46. Kornfeld J-W, Baitzel C, Könner AC, Nicholls HT, Vogt MC, Herrmanns K, et al. Obesity-induced overexpression of miR-802 impairs glucose metabolism through silencing of Hnf1b. Nature. 2013 Feb 7;494(7435):111–5.

47. Verdeguer F, Corre SL, Fischer E, Callens C, Garbay S, Doyen A, et al. A mitotic transcriptional switch in polycystic kidney disease. Nat Med. 2010 Jan;16(1):106–10.

48. Yorifuji T, Kurokawa K, Mamada M, Imai T, Kawai M, Nishi Y, et al. Neonatal Diabetes Mellitus and Neonatal Polycystic, Dysplastic Kidneys: Phenotypically Discordant Recurrence of a Mutation in the Hepatocyte Nuclear Factor-1β Gene Due to Germline Mosaicism. J Clin Endocrinol Metab. 2004 Jun 1;89(6):2905–8.

49. Chen Y-Z, Gao Q, Zhao X-Z, Chen Y-Z, Bennett CL, Xiong X-S, et al. Systematic review of TCF2 anomalies in renal cysts and diabetes syndrome/maturity onset diabetes of the young type 5. Chin Med J (Engl). 2010 Nov;123(22):3326–33.

50. Neve B, Fernandez-Zapico ME, Ashkenazi-Katalan V, Dina C, Hamid YH, Joly E, et al. Role of transcription factor KLF11 and its diabetes-associated gene variants in pancreatic beta cell function. Proc Natl Acad Sci U S A. 2005 Mar 29;102(13):4807–12.

51. Ræder H, Johansson S, Holm PI, Haldorsen IS, Mas E, Sbarra V, et al. Mutations in the CEL VNTR cause a syndrome of diabetes and pancreatic exocrine dysfunction. Nat Genet. 2006 Jan;38(1):54–62.

52. Plengvidhya N, Kooptiwut S, Songtawee N, Doi A, Furuta H, Nishi M, et al. PAX4 mutations in Thais with maturity onset diabetes of the young. J Clin Endocrinol Metab. 2007 Jul;92(7):2821–6.

53. Borowiec M, Liew CW, Thompson R, Boonyasrisawat W, Hu J, Mlynarski WM, et al. Mutations at the BLK locus linked to maturity onset diabetes of the young and β-cell dysfunction. Proc Natl Acad Sci U S A. 2009 Aug 25;106(34):14460–5.

54. Lahiri DK, Nurnberger JI. A rapid non-enzymatic method for the preparation of HMW DNA from blood for RFLP studies. Nucleic Acids Res. 1991 Oct 11;19(19):5444.

55. Rozen S, Skaletsky H. Primer3 on the WWW for general users and for biologist programmers. Methods Mol Biol Clifton NJ. 2000;132:365–86.

56. Schwarz JM, Cooper DN, Schuelke M, Seelow D. MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods. 2014 Apr;11(4):361–2.

57. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, et al. A method and server for predicting damaging missense mutations. Nat Methods. 2010 Apr;7(4):248–9.

58. Choi Y, Sims GE, Murphy S, Miller JR, Chan AP. Predicting the Functional Effect of Amino Acid Substitutions and Indels. PLOS ONE. 2012 out;7(10):e46688.

59. Choi Y. A Fast Computation of Pairwise Sequence Alignment Scores Between a Protein and a Set of Single-locus Variants of Another Protein. In: Proceedings of the ACM Conference on Bioinformatics, Computational Biology and Biomedicine [Internet]. New York, NY, USA: ACM;

2012 [cited 2017 Feb 17]. p. 414–417. (BCB '12). Available from: http://doi.acm.org/10.1145/2382936.2382989

60. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, 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. Genet Med Off J Am Coll Med Genet. 2015 May;17(5):405–24.

61. Wildeman M, van Ophuizen E, den Dunnen JT, Taschner PEM. Improving sequence variant descriptions in mutation databases and literature using the Mutalyzer sequence variation nomenclature checker. Hum Mutat. 2008 Jan;29(1):6–13.

62. MRC-Holland - Products - Details [Internet]. [cited 2017 Jan 29]. Available from: https://www.mlpa.com/WebForms/WebFormProductDetails.aspx?Tag=_tz2fAPIAupKyMjaDF-E-t9bmuxqlhe_Lgqfk8Hkjuss.&ProductOID=_HyxPUmvIJ-w.

63. Hager J, Blanché H, Sun F, Vaxillaire NVM, Poller W, Cohen D, et al. Six Mutations in the Glucokinase Gene Identified in MODY by Using a Nonradioactive Sensitive Screening Technique. Diabetes. 1994 May 1;43(5):730–3.

64. Estalella I, Rica I, De Nanclares GP, Bilbao JR, Vazquez JA, San Pedro JI, et al. Mutations in GCK and HNF-1α explain the majority of cases with clinical diagnosis of MODY in Spain. Clin Endocrinol (Oxf). 2007 outubro;67(4):538–46.

65. Gidh-Jain M, Takeda J, Xu LZ, Lange AJ, Vionnet N, Stoffel M, et al. Glucokinase mutations associated with non-insulin-dependent (type 2) diabetes mellitus have decreased enzymatic activity: implications for structure/function relationships. Proc Natl Acad Sci U S A. 1993 Mar 1;90(5):1932–6.

66. Yellapu NK, Kandlapalli K, Valasani KR, Sarma PVGK, Matcha B. Structural Variations of Human Glucokinase Glu256Lys in MODY2 Condition Using Molecular Dynamics Study. Biotechnol Res Int. 2013 Feb 13;2013:e264793.

67. Yoshiuchi I, Yamagata K, Yang Q, Iwahashi H, Okita K, Yamamoto K, et al. Three new mutations in the hepatocyte nuclear factor-1alpha gene in Japanese subjects with diabetes mellitus: clinical features and functional characterization. Diabetologia. 1999 May;42(5):621–6.

68. Vaxillaire M, Rouard M, Yamagata K, Oda N, Kaisaki PJ, Boriraj VV, et al. Identification of nine novel mutations in the hepatocyte nuclear factor 1 alpha gene associated with maturity-onset diabetes of the young (MODY3). Hum Mol Genet. 1997 Apr;6(4):583–6.

69. Chiu KC, Chuang LM, Ryu JM, Tsai GP, Saad MF. The I27L amino acid polymorphism of hepatic nuclear factor-1alpha is associated with insulin resistance. J Clin Endocrinol Metab. 2000 Jun;85(6):2178–83.

70. Urhammer SA, Fridberg M, Hansen T, Rasmussen SK, Møller AM, Clausen JO, et al. A Prevalent Amino Acid Polymorphism at Codon 98 in the Hepatocyte Nuclear Factor-1α Gene Is Associated With Reduced Serum C-Peptide and Insulin Responses to an Oral Glucose Challenge. Diabetes. 1997 May 1;46(5):912–6.

71. Reiner AP, Gross MD, Carlson CS, Bielinski SJ, Lange LA, Fornage M, et al. Common coding variants of the HNF1A gene are associated with multiple cardiovascular risk phenotypes in community-based samples of younger and older European-American adults: the Coronary Artery Risk Development in Young Adults Study and The Cardiovascular Health Study. Circ Cardiovasc Genet. 2009 Jun;2(3):244–54.

72. Han X, Ji L. [Contribution of MODY2 gene to the pathogenesis of Chinese early onset familial type 2 diabetes]. Beijing Da Xue Xue Bao. 2005 Dec 18;37(6):591–4.

73. Frigeri HR, Martins LT, Auwerter NC, Santos-Weiss ICR dos, Pedrosa FO, de Souza EM, et al. The polymorphism rs2268574 in Glucokinase gene is associated with gestational Diabetes mellitus. Clin Biochem. 2014 abril;47(6):499–500.

74. Giuffrida FMA, Furuzawa GK, Kasamatsu TS, Oliveira MM, Reis AF, Dib SA. HNF1A gene polymorphisms and cardiovascular risk factors in individuals with late-onset autosomal dominant diabetes: a cross-sectional study. Cardiovasc Diabetol. 2009;8:28.

75. Morita K, Saruwatari J, Tanaka T, Oniki K, Kajiwara A, Otake K, et al. Associations between the common HNF1A gene variant p.I27L (rs1169288) and risk of type 2 diabetes mellitus are influenced by weight. Diabetes Metab. 2015 Feb;41(1):91–4.

76. Holmkvist J, Almgren P, Lyssenko V, Lindgren CM, Eriksson K-F, Isomaa B, et al. Common variants in maturity-onset diabetes of the young genes and future risk of type 2 diabetes. Diabetes. 2008 Jun;57(6):1738–44.

77. Wakil SM, Muiya NP, Tahir AI, Al-Najai M, Baz B, Andres E, et al. A New Susceptibility Locus for Myocardial Infarction, Hypertension, Type 2 Diabetes Mellitus, and Dyslipidemia on Chromosome 12q24. Dis Markers [Internet]. 2014 [cited 2016 May 19];2014. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4098619/

78. Bonnycastle LL, Willer CJ, Conneely KN, Jackson AU, Burrill CP, Watanabe RM, et al. Common variants in maturity-onset diabetes of the young genes contribute to risk of type 2 diabetes in Finns. Diabetes. 2006 Sep;55(9):2534–40.

79. Steinthorsdottir V, Thorleifsson G, Sulem P, Helgason H, Grarup N, Sigurdsson A, et al. Identification of low-frequency and rare sequence variants associated with elevated or reduced risk of type 2 diabetes. Nat Genet. 2014 Mar;46(3):294–8.

80. Kawakita R, Hosokawa Y, Fujimaru R, Tamagawa N, Urakami T, Takasawa K, et al. Molecular and clinical characterization of glucokinase maturity-onset diabetes of the young (GCK-MODY) in Japanese patients. Diabet Med. 2014 Nov;31(11):1357–62.

81. Giuffrida FMA, Calliari LE, Manna TD, Ferreira JG, Saddi-Rosa P, Kunii IS, et al. A novel glucokinase deletion (p.Lys32del) and five previously described mutations co-segregate with the phenotype of mild familial hyperglycaemia (MODY2) in Brazilian families. Diabetes Res Clin Pract. 2013 May;100(2):e42–5.

82. Almeida C, Silva SR, Garcia E, Leite AL, Teles A, Campos RA. A novel genetic mutation in a Portuguese family with GCK-MODY. J Pediatr Endocrinol Metab JPEM. 2014 Jan;27(1–2):129–33.

83. de Santana LS, Caetano LA, Costa-Riquetto AD, de Souza Quedas EP, Nery M, Collett-Solberg P, et al. Clinical application of ACMG-AMP guidelines in HNF1A and GCK variants in a cohort of MODY families. Clin Genet. 2017 Feb 1;n/a-n/a.

84. Delvecchio M, Ludovico O, Menzaghi C, Paola RD, Zelante L, Marucci A, et al. Low Prevalence of HNF1A Mutations After Molecular Screening of Multiple MODY Genes in 58 Italian Families Recruited in the Pediatric or Adult Diabetes Clinic From a Single Italian Hospital. Diabetes Care. 2014 Dec 1;37(12):e258–60.

85. Kristinsson SY, Thorolfsdottir ET, Talseth B, Steingrimsson E, Thorsson AV, Helgason T, et al. MODY in Iceland is associated with mutations in HNF-1α and a novel mutation in NeuroD1. Diabetologia. 2001 Nov;44(11):2098–103.

86. Ellard S, Colclough K. Mutations in the genes encoding the transcription factors hepatocyte nuclear factor 1 alpha (HNF1A) and 4 alpha (HNF4A) in maturity-onset diabetes of the young. Hum Mutat. 2006 Sep;27(9):854–69.

87. Boutin P, Gresh L, Cisse A, Hara M, Bell G, Babu S, et al. Missense mutation Gly574Ser in the transcription factor HNF-1alpha is a marker of atypical diabetes mellitus in African-American children. Diabetologia. 1999 Mar;42(3):380–1.

88. Bellanné-Chantelot C, Clauin S, Chauveau D, Collin P, Daumont M, Douillard C, et al. Large Genomic Rearrangements in the Hepatocyte Nuclear Factor-1β (TCF2) Gene Are the Most Frequent Cause of Maturity-Onset Diabetes of the Young Type 5. Diabetes. 2005 Nov 1;54(11):3126–32.

89. Raile K, Klopocki E, Holder M, Wessel T, Galler A, Deiss D, et al. Expanded Clinical Spectrum in Hepatocyte Nuclear Factor 1B-Maturity-Onset Diabetes of the Young. J Clin Endocrinol Metab. 2009 Jul 1;94(7):2658–64.

90. Laffargue F, Bourthoumieu S, Llanas B, Baudouin V, Lahoche A, Morin D, et al. Towards a new point of view on the phenotype of patients with a 17q12 microdeletion syndrome. Arch Dis Child. 2015 Mar;100(3):259–64.

91. Edghill EL, Stals K, Oram RA, Shepherd MH, Hattersley AT, Ellard S. HNF1B deletions in patients with young-onset diabetes but no known renal disease. Diabet Med. 2013 Jan 1;30(1):114–7.

92. Dotto RP, Giuffrida FMA, Franco L, Mathez ALG, Weinert LS, Silveiro SP, et al. Unexpected finding of a whole HNF1B gene deletion during the screening of rare MODY types in a series of Brazilian patients negative for GCK and HNF1A mutations. Diabetes Res Clin Pract. 2016 Jun;116:100–4.

93. Moreno-De-Luca D, Mulle JG, Kaminsky EB, Sanders SJ, Myers SM, Adam MP, et al. Deletion 17q12 Is a Recurrent Copy Number Variant that Confers High Risk of Autism and Schizophrenia. Am J Hum Genet. 2010 Nov 12;87(5):618–30.

94. Palumbo P, Antona V, Palumbo O, Piccione M, Nardello R, Fontana A, et al. Variable phenotype in 17q12 microdeletions: Clinical and molecular characterization of a new case. Gene. 2014 Apr 1;538(2):373–8.

95. Weedon MN, Frayling TM, Shields B, Knight B, Turner T, Metcalf BS, et al. Genetic Regulation of Birth Weight and Fasting Glucose by a Common Polymorphism in the Islet Cell Promoter of the Glucokinase Gene. Diabetes. 2005 Feb 1;54(2):576–81.

96. Weedon MN, Clark VJ, Qian Y, Ben-Shlomo Y, Timpson N, Ebrahim S, et al. A Common Haplotype of the Glucokinase Gene Alters Fasting Glucose and Birth Weight: Association in Six Studies and Population-Genetics Analyses. Am J Hum Genet. 2006 Dec 1;79(6):991–1001.

97. Frigeri HR, Auwerter NC, Koczicki L, Martins LT, de Souza EM, Alberton D, et al. Polymorphisms rs144723656, rs2268574, and rs2268575 of the glucokinase gene are not associated with obese women with type 2 diabetes mellitus. Clin Biochem. 2016 Jan;49(1–2):194–5.

98. Frigeri HR, Santos ICR, Réa RR, Almeida ACR, Fadel-Picheth CMT, Pedrosa FO, et al. Low prevalence of glucokinase gene mutations in gestational diabetic patients with good glycemic control. Genet Mol Res GMR. 2012;11(2):1433–41.

99. Reiner AP, Barber MJ, Guan Y, Ridker PM, Lange LA, Chasman DI, et al. Polymorphisms of the HNF1A Gene Encoding Hepatocyte Nuclear Factor-1α are Associated with C-Reactive Protein. Am J Hum Genet. 2008 May 9;82(5):1193–201.

100. Richter D, Harsch S, Strohmeyer A, Hirobe-Jahn S, Schimmel S, Renner O, et al. MALDI-TOF mass spectrometry screening of cholelithiasis risk markers in the gene of HNF1alpha. J Proteomics. 2012 Jun 27;75(12):3386–99.

101. Li D, Duell EJ, Yu K, Risch HA, Olson SH, Kooperberg C, et al. Pathway analysis of genome-wide association study data highlights pancreatic development genes as susceptibility factors for pancreatic cancer. Carcinogenesis. 2012 Jul;33(7):1384–90.

102. Heikkilä K, Silander K, Salomaa V, Jousilahti P, Koskinen S, Pukkala E, et al. C-reactive protein-associated genetic variants and cancer risk: Findings from FINRISK 1992, FINRISK 1997 and Health 2000 studies. Eur J Cancer. 2011 Feb 1;47(3):404–12.

103. Zhou Y-J, Yin R-X, Hong S-C, Yang Q, Cao X-L, Chen W-X. Association of the HNF1A polymorphisms and serum lipid traits, the risk of coronary artery disease and ischemic stroke. J Gene Med. 2017 Jan;19(1–2).

104. Jafar-Mohammadi B, Groves CJ, Owen KR, Frayling TM, Hattersley AT, McCarthy MI, et al. Low Frequency Variants in the Exons Only Encoding Isoform A of HNF1A Do Not Contribute to Susceptibility to Type 2 Diabetes. PLoS ONE [Internet]. 2009 Aug 12 [cited 2016 May 19];4(8). Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2720540/

105. Liu H-P, Zhao Q, Jin G-Z, Qian Y-W, Gu Y-J, Dong H, et al. Unique genetic alterations and clinicopathological features of hepatocellular adenoma in Chinese population. Pathol - Res Pract. 2015 dezembro;211(12):918–24.

106. Doumatey AP, Chen G, Tekola Ayele F, Zhou J, Erdos M, Shriner D, et al. C-reactive protein (CRP) promoter polymorphisms influence circulating CRP levels in a genome-wide association study of African Americans. Hum Mol Genet. 2012 Jul 1;21(13):3063–72.

107. Pierce BL, Ahsan H. Genome-wide “Pleiotropy Scan” Identifies HNF1A Region as a Novel Pancreatic Cancer Susceptibility Locus. Cancer Res. 2011 Jul 1;71(13):4352–8.

108. Hindorff LA, Sethupathy P, Junkins HA, Ramos EM, Mehta JP, Collins FS, et al. Potential etiologic and functional implications of genome-wide association loci for human diseases and traits. Proc Natl Acad Sci U S A. 2009 Jun 9;106(23):9362–7.

6. Supplementary material

6.1 Comparative analysis of MODY, T1DM and T2DM

Table 6.1: Differentiating MODY from T1DM and T2DM. Multiparametric comparison between GCK-MODY, HNF1A-MODY, HNF4A-MODY, HNF1B-MODY, T1DM and early onset T2DM. DKA: diabetic ketoacidosis. Adapted from (4–6).

	T1DM	Early onset T2DM	GCK-MODY	HNF1A-MODY	HNF4A-MODY	HNF1B-MODY
Insulin dependency	Yes	No	No	No	No	No
Insulin resistance	Rare	Common	Rare	Rare	Rare	Rare[2]
Parental family history	2-4%	Typically	Yes	Yes	Yes	Yes
Age of onset	6 months to young adult	Teen to young adult	Birth	Teen to young adult	Teen to young adult	Teen to young adult
Polyuria/ polydipsia	Common	Variable	Variable	Variable	Variable	Variable
Obesity	Population frequency	Yes	Population frequency	Population frequency	Population frequency	Population frequency
Acanthosis nigricans	No	Yes	No	No	No	No
DKA	Common	Rare	Rare	Rare	Rare	Rare
Microvascular complications	Common	Common	Rare	Common	Common	Common
Other features	None	None	None	Glycosuria Normal/elevated HDL	Macrosomia Low HDL & triglycerides Elevated LDL Neonatal hyperinsulinemia	Genital & renal anomalies Renal cysts
Treatment	Insulin	Exercise OHA progressing to insulin	Not needed or diet	Sulphonylurea May progress to insulin	Sulphonylurea May progress to insulin	Insulin
FPG (mmol/L)	Often >7.1	Often >7.1	>5.5; mild	Often <5.5	Often <5.5	Often<5.5
OGTT increment (mmol/L)	Typically >3.5	Typically >3.5	Typically <3.0	Typically >3.5	Typically >3.5	Typically >3.5
Serum C-peptide (ρmol/L)	<200 (>5 years)	500->1000	100-900	100-700	100-700	100-700
Pancreatic autoantibodies	80-90% (at diagnosis)	1-20%	<1% (same as controls)	<1% (same as controls)	<1% (same as controls)	<1% (same as controls)
Pathophysiology	Autoimmune β-cell destruction	Insulin resistance and/or relative β-cell deficiency	Glucose sensing defect in β-cell	Insulin secretion defect in β-cell	Insulin secretion defect in β-cell	Insulin secretion defect in β-cell

[2] About 50% *HNF1B* diabetic mutation carriers have β-cell dysfunction, relative to insulin secretion, and insulin resistance (4).

6.2 Online tools

- Primer3Plus: http://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi
- SNPCheck: https://secure.ngrl.org.uk/SNPCheck/snpcheck.htm
- UCSC *in silico* PCR: https://genome.ucsc.edu/cgi-bin/hgPcr?hgsid=586569735_R9ktENtkdSaprva4TQoaAauWGzM4
- UCSC BLAT: https://genome.ucsc.edu/cgi-bin/hgBlat?hgsid=586569735_R9ktENtkdSaprva4TQoaAauWGzM4&command=start
- MutationTaster: http://www.mutationtaster.org/
- PolyPhen-2: http://genetics.bwh.harvard.edu/pph2/
- PROVEAN/SIFT: http://provean.jcvi.org/genome_submit_2.php

6.3 Interpreting *in silico* mutation predictions

MutationTaster gives "disease causing" or "polymorphism" (*i.e.* benign) predictions, coupled with a probability value. A value close to 1 indicates a high "security" prediction. Probability values under 0.5 indicate the automatic prediction for a variant differs from the classification MutationTaster would have made[3]. PolyPhen-2 gives "benign", "possibly damaging" or "probably damaging" predictions, with the score scale going from 0 (benign) to 1 (probably damaging). PROVEAN gives "neutral" (> -2.5) or "deleterious" (≤ -2.5) predictions; and SIFT gives "tolerated" (>0.05) or "damaging" (<0.05) predictions.

6.4 Primer sequences, PCR master mix and cycling program

6.4.1 Primer sequences

Table 6.2: Primer data. Data on primer sequence, size (bp), annealing temperature (ºC), orientation (forward or reverse), amplicon size (bp) and genomic location.

Primer sequence (5'-3')	Size (bp)	Annealing temperature (ºC)	Primer type	Amplicon size (bp)	Genomic location	Gene
CCTGTCCAGCTTTGGACTCT	20	62	F	598	Promoter	***GCK* (MODY2)**
GTGACCTGGGGACAGCTTTT	20		R			
TGCATGGCAGCTCTAATGAC	20	62	F	754	Exon 1	
CCTTCTCAAAGAGCCTGTGC	20		R			
GTGTGCAGATGCCTGGTG	18	62	F	346	Exon 2	
CTGGCTGTGAGTCTGGGAGT	20		R			
CCTTCCCTCCTCCTCTTTGT	20	61	F	364	Exon 3	
CCACCCCTGGTAGACAGGT	19		R			
CGGAAGAGGAGAGGGAAACT	20	61; 62	F	382	Exon 4	
CCAGATCTCCCTTCTGAGCA	20		R			
CTGCTCTGAGCCTGTTTCCT	20	62	F	565	Exons 5 and 6	
ACCAGGCTCTGCTCTGACAT	20		R			
CCATTGTTCCAGACAAAGCA	20	60; 61	F	400	Exon 7	
CAAGCCCATTATCTGCAATG	20		R			
GACATTTCTAAAGCTCTGGCTCA	23	60	F	698	Exon 8	
CATCCCTGCTCTTGGCATC	19		R			

[3] For more information see http://www.mutationtaster.org/info/FAQs.html

GCTCAGCGAGGGAAAGAG	18	60; 61	F	482	Exon 9	
GGGGGACGAGAAGAGGACTA	20		R			
GTCCCCCTTGGCCTAGATT	19	61; 62	F	698	Exon 10a[4]	
AGGTCTGGTCAAGCTGTTGG	20		R			
CTACAGGGAGGCCTGGTGTC	20	61	F	585	Promoter	***HNF1A* (MODY3)**
CGGCAGACACAAACCAAACT	20		R			
AGTCCCTTCGCTAAGCACAC	20	59-**61**	F	671	Exon 1	
GGCTCGTTAGGAGCTGAGG	19		R			
CCCTGTGTCCTGGCATAAAT	20	**59**-61	F	489	Exon 2	
TGTGTAATGGGGATGGTGAA	20		R			
GCATGTGTGCTGTGTGTTTG	20	61	F	491	Exon 3	
GCCAGGCTAAGCCAATATCA	20		R			
CAGATCTGCCAGCCTCAAAC	20	58	F	480	Exon 4	
CATGAATGGAATGGAACCAA	20		R			
GCCTAAGCAAACCAATGGAG	20	61	F	655	Exons 5 and 6	
CTCTTCCAGCTCCTGGATTC	20		R			
CTCTGGGAAGGAGAGGTGGT	20	60; **61**	F	397	Exon 7	
GTCCCAGAGACACATGCAGA	20		R			
AGTCTTGAGGCCTGGGACTA	20	59; 60; 63	F	558	Exons 8 and 9	
CTTCCTCACAGCAGCCCTA	19		R			
GGTGTGACTTTGGGGTTCC	19	61	F	850	Exon 10a[5]	
CAGAGTAGCCACCCAGGAAA	20		R			

F: forward primer. R: reverse primer.

6.4.2 PCR master mix and cycling program

Table 6.3: Standard PCR mix and cycling program. Minor adjustments in annealing temperature, cycle number or $MgCl_2$ volume were made to optimize some fragments' amplification. Each PCR reaction required 1 µL DNA sample (typically at 100 ng/µL) added to 24 µL master mix.

Cycling program			PCR Mix (for one reaction)	
			Reagent	Vol (µL): 1x
Temperature (ºC)	Time		H20	14.625
95	5 min		dNTP's	4
94	45 sec	38 cycles	Buffer	2.5
Annealing	45 sec		$MgCl_2$	0.75
72	60 sec		F Primer	1
72	7 min		R Primer	1
4	15 min/∞		DNA polymerase	0.125
			Total volume	24

[4] Primer pair covers terminating coding region but does not cover the entire exon 10.
[5] Primer pair covers terminating coding region but does not cover the entire exon 10.

6.5 Sanger sequencing mix and cycling program

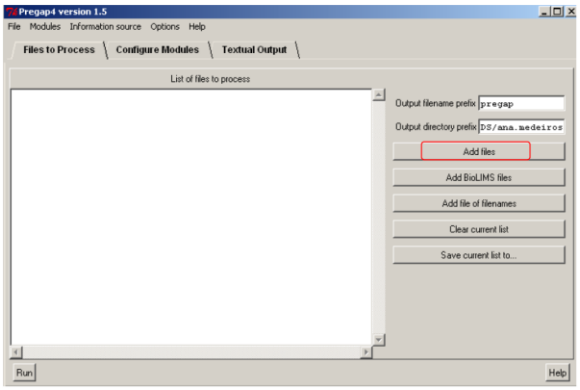
Table 6.4: Sequencing mix and cycling program. Typical sequencing mix (right) and cycling program (left). BigDye volume was adjusted considering fragment size. 1 µL was used for fragments under 400 bp; 1.5 µL for fragments up to 500 bp; and 2 µL for fragments above 500 bp. Purified PCR product volume was adjusted considering DNA sample concentration. For concentrations under 100 ng/µL, 2 µL purified PCR product were used; and above 100 ng/µL, 1 µL purified PCR product was used.

Temperature (ºC)	Time		Reagents	Volume (µL)
96	1 min		Sterile H_2O	6
96	10 sec	25 cycles	Primer (2 ρmol/µL)	1
58	5 sec		BigDye reaction mix	2
55	4 min		Purified PCR product	0.5-2
4	∞			

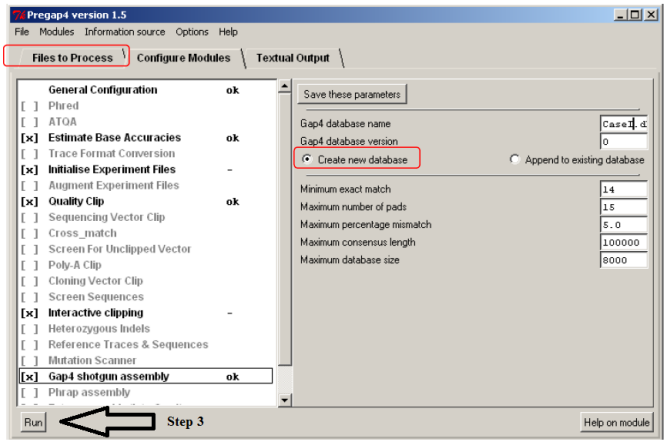
6.6 Sequence analysis and pregap module configuration

Staden Package chromatogram analysis was done by aligning sequence files/chromatograms and, in Gap4, manually screening for variants against a reference sequence and other subjects' sequence files/chromatograms. Before this, Pregap4 was used to assemble sequence files into contigs, following these steps:

1. Add desired sequence files to the list of files to process;



2. Set Pregap4 module configuration as illustrated;



3. Hit run to assemble contig;
4. Open the created database in Gap4 to screen for variants as previously described.

6.7 Detected sequence variants

Table 6.5: Detected sequence variants. All *GCK*, *HNF1A* and *HNF1B* detected sequence variants are listed below. MAF's taken from each sequence variant dbSNP entry regard European (EUR) and African (AFR) populations.

***GCK* (NG_008847.1; NM_000162.3; NP_000153.1)**								
dbSNP	**Alteration (genomic)**	**Alteration (transcript)**	**Location**	**Alteration (protein)**	**MAF (EUR)**	**MAF (AFR)**	**ACMG classification**	**References**
rs35670475	g.4703C>T	c.-768C>T	Before promoter	p.(=)	A=0.06460001	A=0.00450000	Benign	-
rs12702070	g.4791A>G	c.-680A>G	Before promoter	p.(=)	C=0.06460001	C=0.00830000	Benign	(16)
rs1799884	g.4955G>A	c.-516G>A	Putative promoter	p.(=)	A=0.17790000	A=0.16410001	Benign	(16,76,95)
rs13306391	g.5387C>G	c.-84C>G	5' UTR	p.(=)	G=0.06660000	G=0.06730000	Benign	(16,96)
-	g.43349C>T	c.364C>T	Exon 4	p.(Leu122Phe)	-	-	Likely pathogenic	(16,80)
rs115206967	g.43526A>T	c.483+58A>T	Intron 4	p.(=)	A=0.00100000	A=0.21630001	Benign	(16)
rs2268573	g.43555A>C	c.483+87A>C	Intron 4	p.(=)	A=0.48510000	A=0.20880000	Benign	(16)
-	g.(43915_48326)del	c.484-?_1019+?del	Exons 5-8	p.(Gly162Argfs*118)	-	-	Pathogenic	-
-	g.44456_44488del	c.579+1_579+33del	Intron 5	p.(?)	-	-	Pathogenic	(16,63,81)
rs2268574	g.44702T>C	c.679+38T>C	Intron 6	p.(=)	T=0.48510000	T=0.22010000	Benign	(16,73)
rs2268575	g.44749A>G	c.679+85A>G	Intron 6	p.(=)	G=0.19479999	G=0.16410001	Benign	(97)
rs769268803	g.46677G>A	c.766G>A	Exon 7	p.Glu256Lys	-	-	Pathogenic	(16,64–66)
rs76323047	g.48068T>C	c.1019+107T>C	Intron 8	p.(=)	G=0.11930000	G=0.05300000	Benign	-
rs887686	g.48135G>T	c.1019+174G>T	Intron 8	p.(=)	T=0.29319999	T=0.02270000	Benign	-
rs887687	g.48218G>C	c.1019+257G>C	Intron 8	p.(=)	C=0.29220000	C=0.03780000	Benign	-
rs5883890	g.48236del	c.1019+275del	Intron 8	p.(=)	-=0.34489998	-= 0.12860000	Benign	-
rs2971680	g.48265G>C	c.1019+304G>C	Intron 8	p.(=)	C=0.05770000	G=0.37590000	Benign	-
rs2908274	g.48935C>T	c.1253+8C>T	Intron 9	p.(=)	T=0.17690000	C=0.27460000	Benign	(16,72,98)
rs193922273	g.49158T>A	c.1268T>A	Exon 10a	p.(Phe423Tyr)	-	-	Likely pathogenic	(16)
rs13306388	g.49620G>A	c.*332G>A	3' UTR	p.(=)	A=0.26539999	A=0.02190000	Benign	(98)
***HNF1A* (NG_011731.2; NM_000545.6; NP_000536.5)**								
dbSNP	**Alteration (genomic)**	**Alteration (transcript)**	**Location**	**Alteration (protein)**	**MAF (EUR)**	**MAF (AFR)**	**ACMG classification**	**References**

rs1169289	g.5074C>G	c.51C>G	Exon 1	p.(=)	C=0.46919999	C=0.31690001	Benign	-
rs1169288	g.5102A>C	c.79A>C	Exon 1	p.(Ile27Leu)	G=0.33899999	G=0.08320000	Benign	(69,71,74–76)
rs34056805	g.5299C>T	c.276C>T	Exon 1	p.(=)	T=0.00000000	T=0.01210000	Likely benign	-
rs1800574	g.5316C>T	c.293C>T	Exon 1	p.(Ala98Val)	T=0.02780000	T=0.00300000	Likely benign	(70,75)
rs1169293	g.14997G>A	c.327-91G>A	Intron 1	p.(=)	G=0.08550000	G=0.21709999	Benign	(99)
rs1169294	g.15046G>A	c.327-42G>A	Intron 1	p.(=)	A=0.33600000	A=0.14980000	Benign	(100)
rs12427353	g.15353G>C	c.526+66G>C	Intron 2	p.(=)	C=0.20080000	C=0.01130000	Benign	-
rs1169300	g.19677G>A	c.527-98G>A	Intron 2	p.(=)	A=0.31610000	A=0.11800000	Benign	(101,102)
rs2071190	g.19724T>A	c.527-51T>A	Intron 2	p.(=)	A=0.23559999	A=0.23600000	Benign	(78)
rs1169301	g.19752C>T	c.527-23C>T	Intron 2	p.(=)	T=0.31610000	T=0.11800000	Benign	-
-	g.20519C>T	c.814C>T	Exon 4	p.(Arg272Cys)	-	-	Pathogenic	(67)
rs56348580	g.20569G>C	c.864G>C	Exon 4	p.(=)	C=0.30320001	C=0.17780000	Benign	(79)
-	g.20577del	c.872del	Exon 4	p.(Pro291Glnfs*51)	-	-	Pathogenic	(11,68)
rs55783344	g.20751C>T	c.955+91C>T	Intron 4	p.(=)	T=0.02290000	T=0.00230000	Insufficient data	-
rs1169302	g.20754T>G	c.955+94T>G	Intron 4	p.(=)	G=0.44530001	G=0.27160001	Benign	-
rs3751156	g.22754G>T	c.1108-42G>T	Intron 5	p.(=)	A=0.02290000	A=0.00230000	Insufficient data	-
rs138835108	g.22769C>T	c.1108-27C>T	Intron 5	p.(=)	T=0.00300000	T=0.00080000	Insufficient data	-
rs2259820	g.23794C>T	c.1375C>T	Exon 7	p.(=)	A=0.31610000	A=0.10890000	Benign	(77,103)
rs2464196	g.23879G>A	c.1460G>A	Exon7	p.(Ser487Asn)	T=0.31610000	T=0.11200000	Benign	(71,74,77)
rs2464195	g.23927G>A	c.1501+7G>A	Intron 7	p.(=)	T=0.37770000	T=0.12629999	Benign	(99)
rs55834942	g.25566G>A	c.1545G>A	Exon 8	p.(=)	A=0.19880000	A=0.01210000	Benign	(104)
rs1169304	g.25673T>C	c.1623+29T>C	Intron 8	p.(=)	A=0.16600001	A=0.45989999	Benign	(104,105)
rs1169305	g.25834G>A	c.1720G>A	Exon 9	p.(Gly574Ser)	T=0.00000000	T=0.05370000	Uncertain significance	(87)
rs76003406	g.25926C>T	c.1768+44C>T	Intron 9	p.(=)	T=0.05570000	T=0.00150000	Benign	-
rs735396	g.27296T>C	c.1769-24T>C	Intron 9	p.(=)	G=0.37770000	G=0.12629999	Benign	(101,104)
rs1169309	g.27644G>T	c.*197G>T	3' UTR	p.(=)	T=0.37670001	T=0.12629999	Benign	-
rs1169310	g.27885G>A	c.*438G>A	3' UTR	p.(=)	A=0.37770000	A=0.12629999	Benign	(106–108)
***HNF1B* (NG_013019.2; NM_000458.3; NP_000449.1)**								
dbSNP	**Alteration (genomic)**	**Alteration (transcript)**	**Location**	**Alteration (protein)**	**MAF (EUR)**	**MAF (AFR)**	**ACMG classification**	**References**
-	-	c.1-?_1674+?del	Exons 1-9	p.(?)	-	-	Pathogenic	(11,88,89,91,92)