Maturity-onset diabetes of the young type 5 (MODY5) in a family with diabetes and mild kidney disease diagnosed by whole exome sequencing.

John M Wentworth[1,2,4], Vesna Lukic[1], Melanie Bahlo[1,4,5], Moira Finlay[3], Cao Nguyen[6,7], Grant Morahan[6,7], Leonard C Harrison[1,4]

1. The Walter and Eliza Hall Institute of Medical Research, Parkville 3052, Australia
2. Department of Diabetes and Endocrinology, Royal Melbourne Hospital, Parkville 3050, Australia
3. Department of Anatomical Pathology, Royal Melbourne Hospital, Parkville 3050, Australia
4. University of Melbourne, Department of Medical Biology, Melbourne 3010, Australia
5. University of Melbourne, Department of Mathematics and Statistics, Melbourne 3010, Australia
6. Centre for Diabetes Research, The Western Australian Institute for Medical Research, Australia
7. Centre for Medical Research, University of Western Australia, Australia

Corresponding Author:
John M Wentworth
Walter and Eliza Hall Institute
1G Royal Parade
Parkville, 3052
Australia
wentworth@wehi.edu.au
Ph +61 3 93452555
Fax +61 3 93470852

Keywords: MODY5, HNF1B, glomerulocystic kidney disease, maturity-onset diabetes of the young, whole exome sequencing

Word count: 1370

In 1967, the mother, aged 16 years and not overweight, presented with thirst and polyuria following a bout of tonsillitis. Hyperglycaemia was confirmed and she commenced insulin therapy for presumed type 1 diabetes. At the same time, she was noted to have moderate renal impairment (creatinine 0.17mM), normal urine microscopy and <1g/24h proteinuria. Her pregnancy at age 22 years was complicated by hypertension and further deterioration of renal function. Her male child was delivered at 33 weeks gestation weighing 1260g, 40g below the 3rd centile for age. In 1975, she had a renal biopsy that was reported to show minor, non-specific features that included focal tubular atrophy, interstitial fibrosis and mesangial proliferation. A second pregnancy to the same father delivered another son, of normal birth weight (2130g), at 33 weeks gestation. Her renal function then progressively deteriorated, and she required dialysis at age 50 followed by combined kidney/pancreas transplantation at age 53. At the time of her transplant, her fasting C-peptide level was below the limit of detection of 0.166nmol/l. Her medical history also included low birth weight (self-reported as 4lb or 1814g at 38 weeks gestation) and rheumatoid arthritis diagnosed at age 52 years.

The first son, who had also never been overweight, was diagnosed with type 1 diabetes at age 17 years when he developed thirst and polyuria associated with non-ketotic hyperglycaemia. He had an elevated serum creatinine (0.17mM), mild proteinuria (0.26g/24h) and hypertension (blood pressure 150/100mmHg). Renal ultrasound showed mildly hyperechogenic kidneys of normal size (each 9.5cm), with two small cysts on the left kidney. A renal biopsy was reported as showing non-specific findings of mild mesangial thickening and tubular fibrosis. The similarities of his condition to that of his mother prompted discussion between his nephrologist and the renal histopathologist, but they could not arrive at a unifying diagnosis. His diabetes remained well controlled (HbA1c<6.2%) despite varying compliance with insulin therapy. Multiple tests for islet autoantibodies were negative and fasting C-peptide five years after diabetes diagnosis was normal at 0.68nmol/l. At age 38, he and his partner were assessed for infertility. Semen analysis showed 17 million sperm per milliliter (RR>15 million/ml), of which 98% were dysmorphic (RR<96%) and 33% were motile (RR>40%).

The extended family tree is shown in Figure 1. The second son had not been diagnosed with either diabetes or kidney disease at age 38. The father did not have diabetes or kidney disease as a young adult, but his health in older age is not known because he is estranged. On the mother's side of the family, her sister developed rheumatoid arthritis aged 21 years, one brother became obese in early adulthood and developed diabetes aged 50 years and the other brother died in his 20s in a motor vehicle accident. Her father was thought to be overweight but not obese and developed diabetes aged 70 years whereas her mother developed rheumatoid arthritis aged 50 years and diabetes, thought to be related to steroid therapy, at 52 years of age.

We hypothesized that a rare gene mutation present in the mother and her affected son accounted for their syndrome of young-onset diabetes and kidney disease. We therefore sought permission from Melbourne Health Human Research Ethics Committee to have the Australian Genome Research Facility perform whole exome sequencing on whole blood DNA isolated from the mother and both sons. We obtained 10-fold or better genome coverage for 96% of the exomes and used our analysis pipeline (1) to identify non-synonymous base changes whose allele frequencies within the Australian population were less than 1%. We identified 50 non-synonymous base changes that occurred only in the affected mother and son and that were

predicted by the SIFT and Polyphen algorithms (2) to impair protein function (Supplementary table). Only one of these mutations occurred at a locus previously associated with diabetes and/or kidney disease: an A-G base change in the *HNF1B* gene predicting a M160V amino acid substitution in a highly conserved mutation hotspot within a DNA-binding domain (3). Heterozygosity for this change was confirmed in the mother and affected son by direct sequencing.

Methionine and valine are both non-polar amino acids with similar steric properties, suggesting the M160V substitution might not impair HNF1B protein function. However, this variant has been previously identified in a kindred with hyperechogenic kidneys, renal cysts and glomerulocystic changes in a renal biopsy (4). This suggested that our kindred's *HNF1B* mutation was pathogenic and prompted us to look for similar changes in the affected son's biopsy (the mother's specimen had been destroyed). Although not described in the original histopathology report, 6 of 18 glomeruli had markedly dilated Bowman's spaces indicative of glomerulocystic disease (Figure 2).

We also imputed HLA-DR genotype by genotyping the mother and her sons for three single nucleotide polymorphisms (5). The mother and her unaffected son carried the high-risk type 1 diabetes genotype (HLA-DR3 combined with HLA-DR4 alleles) whereas the affected son carried an intermediate risk genotype (one HLA-DR4 allele only). Taken with the abovementioned C-peptide and islet autoantibody findings, these HLA-DR genotypes are consistent with a non-autoimmune cause of diabetes affecting the son but leave open the possibility of autoimmune type 1 diabetes affecting the mother.

We concluded that several seemingly disparate clinical features in the mother and her affected son represented the MODY5 syndrome, caused by a M160V amino acid change in HNF1B. These features were: young-onset diabetes affecting the mother and son (6), hyperechogenic cystic kidneys noted in the son's ultrasound (4, 7, 8), the low birth weight of the mother and son (9) and the son's infertility, which may have reflected genital tract malformation (8). This is the first reported case of MODY5 associated with this *HNF1B* mutation.

Maturity onset diabetes of the young (MODY) is a rare monogenic syndrome of non-ketotic diabetes that affects successive generations of a family before 25 years of age (10). The three most common MODY subtypes (MODY1, -2 and -3) are associated with mutations in genes critical for pancreatic beta cell development (*HNF1A* and *HNF4A*) and function (*GCK*). The MODY5 subtype is uniquely associated with glomerulocystic kidney disease and is due to *HNF1B* gene mutation. More than 30 different *HNF1B* mutations have been described, but most have been found in families affected by glomerulocystic kidney disease without diabetes (3, 11). Most *HNF1B* gene mutations localise to one of two DNA-binding domains, which are postulated to target the HNF1B protein to promoters of genes critical for pancreatic and renal development. The M160V mutation observed in our family has previously been identified in a kindred with glomerulocystic kidney disease without diabetes (4).

The exome comprises the regions of the genome that encode proteins, accounting for around 1% of DNA in humans. Exome sequencing involves amplifying these protein-coding regions by polymerase chain reaction and sequencing around a million of these simultaneously. The

sequences are then aligned and cross-referenced to the human reference genome to identify changes in the genetic code. Although exome sequencing is a powerful and cost-effective method for mutation screening, it may not provide reliable sequence for around 5% of exomes, and the large amounts of data demand stringent bioinformatic approaches to verify data quality and to correctly interpret the many thousands of genetic polymorphisms (12). In addition, a mutation cannot be ascribed to an individual before it is confirmed by traditional gene sequencing.

Exome sequencing is currently the most widely used technology for the identification of coding changes in the genome and was, until recently, much cheaper than whole genome sequencing. However, the recent advent of the '$1000 genome' (13) will make exome sequencing virtually redundant. Whole genome sequencing enables more uniform and comprehensive sequencing coverage, translating into more variants captured (14). The interpretation of these variants, particularly those occurring in non-coding (non-exomic) DNA, will be challenging.

This case illustrates how exome sequencing of small kindreds can help refine clinical diagnosis. In this instance, identification of a *HNF1B* mutation provided the affected mother and son with an explanation for several of their health problems and enabled disease screening of other family members. It also highlights the importance of a multidisciplinary approach to interpreting the vast amount of genetic variability identified by whole exome sequencing, especially relevant given its rapid uptake in the clinical arena (15).

Acknowledgements
The authors thank the family for participating in this study. This work was funded by program (1037321) and infrastructure grants from the National Health and Medical Research Council of Australia (NHMRC), a Victorian State Government Operational Infrastructure Support Grant, the Diabetes Australia Research Trust, and the Royal Australian College of Physicians Research Foundation.

References

1. Smith, K.R., Bromhead, C.J., Hildebrand, M.S., Shearer, A.E., Lockhart, P.J., Najmabadi, H., Leventer, R.J., McGillivray, G., Amor, D.J., Smith, R.J., et al. 2011. Reducing the exome search space for mendelian diseases using genetic linkage analysis of exome genotypes. *Genome Biol* 12:R85.
2. Ng, P.C., and Henikoff, S. 2006. Predicting the effects of amino acid substitutions on protein function. *Annu Rev Genomics Hum Genet* 7:61-80.
3. Lu, P., Rha, G.B., and Chi, Y.I. 2007. Structural basis of disease-causing mutations in hepatocyte nuclear factor 1beta. *Biochemistry* 46:12071-12080.
4. Bergmann, C., von Bothmer, J., Ortiz Bruchle, N., Venghaus, A., Frank, V., Fehrenbach, H., Hampel, T., Pape, L., Buske, A., Jonsson, J., et al. 2011. Mutations in multiple PKD genes may explain early and severe polycystic kidney disease. *J Am Soc Nephrol* 22:2047-2056.
5. Nguyen, C., Varney, M.D., Harrison, L.C., and Morahan, G. 2013. Definition of high-risk type 1 diabetes HLA-DR and HLA-DQ types using only three single nucleotide polymorphisms. *Diabetes* 62:2135-2140.
6. Stride, A., and Hattersley, A.T. 2002. Different genes, different diabetes: lessons from maturity-onset diabetes of the young. *Ann Med* 34:207-216.
7. Bellanne-Chantelot, C., Chauveau, D., Gautier, J.F., Dubois-Laforgue, D., Clauin, S., Beaufils, S., Wilhelm, J.M., Boitard, C., Noel, L.H., Velho, G., et al. 2004. Clinical spectrum associated with hepatocyte nuclear factor-1beta mutations. *Ann Intern Med* 140:510-517.
8. Faguer, S., Decramer, S., Chassaing, N., Bellanne-Chantelot, C., Calvas, P., Beaufils, S., Bessenay, L., Lengele, J.P., Dahan, K., Ronco, P., et al. 2011. Diagnosis, management, and prognosis of HNF1B nephropathy in adulthood. *Kidney Int* 80:768-776.
9. Edghill, E.L., Bingham, C., Slingerland, A.S., Minton, J.A., Noordam, C., Ellard, S., and Hattersley, A.T. 2006. Hepatocyte nuclear factor-1 beta mutations cause neonatal diabetes and intrauterine growth retardation: support for a critical role of HNF-1beta in human pancreatic development. *Diabet Med* 23:1301-1306.
10. Fajans, S.S., Bell, G.I., and Polonsky, K.S. 2001. Molecular mechanisms and clinical pathophysiology of maturity-onset diabetes of the young. *N Engl J Med* 345:971-980.
11. Heidet, L., Decramer, S., Pawtowski, A., Moriniere, V., Bandin, F., Knebelmann, B., Lebre, A.S., Faguer, S., Guigonis, V., Antignac, C., et al. 2010. Spectrum of HNF1B mutations in a large cohort of patients who harbor renal diseases. *Clin J Am Soc Nephrol* 5:1079-1090.
12. MacArthur, D.G., Manolio, T.A., Dimmock, D.P., Rehm, H.L., Shendure, J., Abecasis, G.R., Adams, D.R., Altman, R.B., Antonarakis, S.E., Ashley, E.A., et al. 2014. Guidelines for investigating causality of sequence variants in human disease. *Nature* 508:469-476.
13. Hayden, E.C. 2014. Technology: The $1,000 genome. *Nature* 507:294-295.

14. Shendure, J. 2011. Next-generation human genetics. *Genome Biol* 12:408.
15. Solomon, B.D. 2014. Incidentalomas in genomics and radiology. *N Engl J Med* 370:988-990.

Figure 1. Family tree. Shaded symbols represent diabetes and renal impairment in the mother and her affected son. Family members who provided DNA samples are indicated by *. RA: rheumatoid arthritis; MVA: motor vehicle accident.

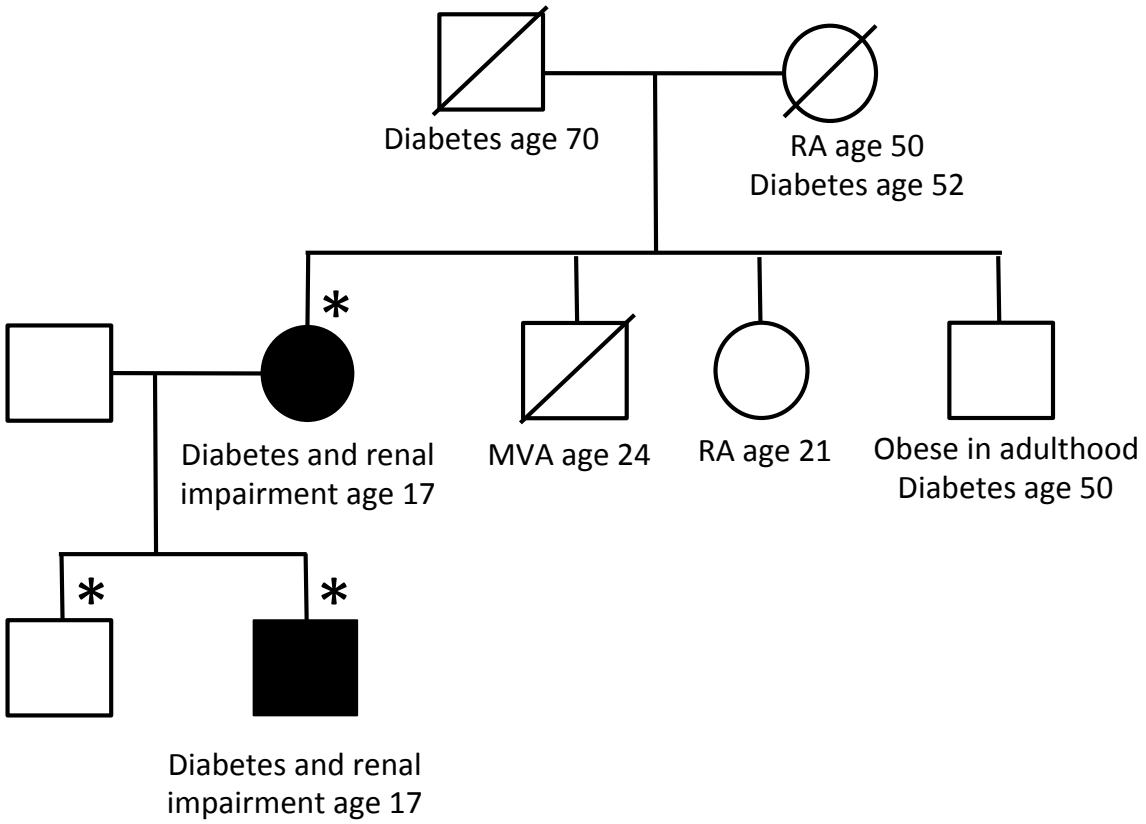
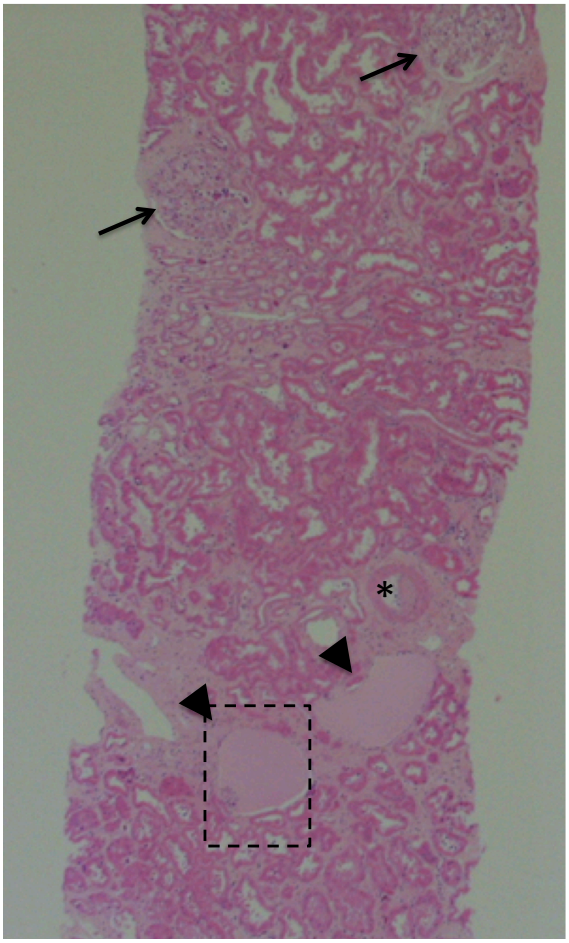
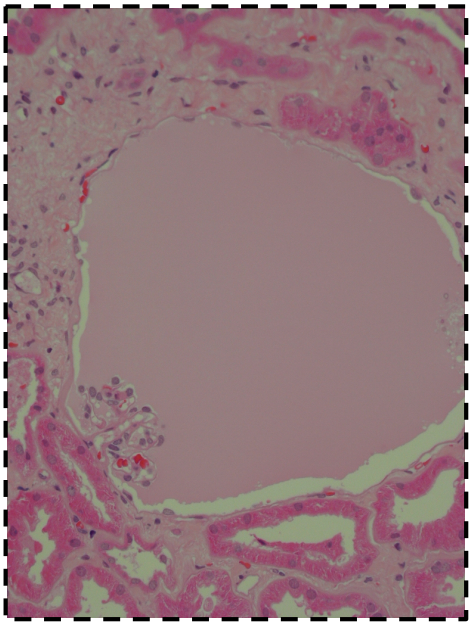


Figure 2. **Glomerulocystic change in the affected son's kidney.**
The arrows and arrowheads in the low-power (x4) view indicate functional glomeruli and glomerular cysts respectively. The asterix indicates a small artery showing moderate arteriosclerosis with surrounding fibrosis. The high-power (x40) view of the boxed area shows a glomerular tuft within a dilated Bowman's space, indicative of glomerulocystic disease.





Supplementary Table
Details of the 50 non-synonymous mutations that met our selection criteria

Symbol	Gene	UniProt Accession No.	Chromosome	Base position	Reference	Mutation	Amino acid substitution	Allele frequency
ARHGAP42	Rho GTPase activating protein 42	A6NI28	11	100830655	A	G	I402V	novel
ARL5B	ADP-ribosylation factor-like 5B	Q96KC2	10	18955518	A	T	I21L	novel
ATOH8	atonal homolog 8	Q96SQ7-2	2	85981734	A	C	E141A	novel
ATP8A2	ATPase, aminophospholipid transporter, class I, type 8A, member 2	Q9NTI2	13	26153950	G	T	L584F	novel
ATP8B4	ATPase, class I, type 8B, member 4	Q8TF62	15	50264839	C	T	G395S	0.005931
C17orf57	chromosome 17 open reading frame 57	Q8IY85	17	45452309	A	G	E450G	novel
C1orf96	chromosome 1 open reading frame 96	Q6IQ19	1	229461083	C	G	A238P	0.001
CHI3L2	chitinase 3-like 2	Q15782	1	111783982	C	T	R318W	novel
CLEC18A	C-type lectin domain family 18, member A	A5D8T8	16	69996928	G	A	D421N	novel
CLECL1	C-type lectin-like 1	Q8IZS7	12	9875323	T	C	T135A	0.007612
CNRIP1	cannabinoid receptor interacting protein 1	Q96F85	2	68521098	G	A	R131W	novel
CPT1C	carnitine palmitoyltransferase 1C	Q8TCG5-3	19	50210874	G	A	M481I	novel
CSMD1	CUB and Sushi multiple domains 1	E5RIG2	8	3087687	C	G	R1408T	0.005088
DAGLB	diacylglycerol lipase, beta	B4DQU0	7	6476161	G	C	T84R	0.003152
DCAF5	DDB1 and CUL4 associated factor 5	G3V4J7	14	69521652	C	T	R583Q	0.001076
DIS3	DIS3 mitotic control homolog	Q9Y2L1-2	13	73347859	G	A	A371V	0.001846
EPN3	epsin 3	F6QWW5	17	48615458	G	T	R249L	0.001384
EVC2	Ellis van Creveld syndrome 2	Q86UK5	4	5620263	G	A	A883V	0.002691
FAT2	FAT tumor suppressor homolog 2	Q9NYQ8	5	150924344	T	C	Y2115C	0.000154
FBN3	fibrillin 3	Q75N90	19	8152724	C	T	G2202R	0.001307
FNDC1	fibronectin type III domain containing 1	Q4ZHG4-2	6	159660915	C	A	T1453N	0.001259
HECW1	HECT, C2 and WW domain containing E3 ubiquitin protein ligase 1	Q76N89	7	43540841	T	C	V1184A	0.003321
HHIPL2	HHIP-like 2	Q6UWX4	1	222696035	G	A	R695C	0.001615