

Exome Sequencing and Genetic Testing for MODY

Stefan Johansson^{1,2,3}, Henrik Irgens^{1,4}, Kishan K. Chudasama^{1,2}, Janne Molnes¹, Jan Aerts^{5,6}, Francisco S. Roque⁷, Inge Jonassen^{7,8}, Shawn Levy⁹, Kari Lima¹⁰, Per M. Knappskog^{1,2}, Graeme I. Bell¹¹, Anders Molven^{12,13}, Pål R. Njølstad^{1,4*}

1 Department of Clinical Medicine, University of Bergen, Bergen, Norway, **2** Center for Medical Genetics and Molecular Medicine, Haukeland University Hospital, Bergen, Norway, **3** Department of Biomedicine, University of Bergen, Bergen, Norway, **4** Department of Pediatrics, Haukeland University Hospital, Bergen, Norway, **5** Faculty of Engineering – ESAT/SCD, Leuven University, Leuven, Belgium, **6** Wellcome Trust Sanger Institute, Cambridge, United Kingdom, **7** Computational Biology Unit, Uni Computing, Uni Research, Bergen, Norway, **8** Department of Informatics, University of Bergen, Bergen, Norway, **9** HudsonAlpha Institute for Biotechnology, Huntsville, Alabama, United States of America, **10** Division of Medicine, Department of Endocrinology, Departments of Medicine and Human Genetics, Akershus University Hospital, Lørenskog, Norway, **11** Departments of Medicine and Human Genetics, The University of Chicago, Chicago, Illinois, United States of America, **12** Gade Institute, University of Bergen, Bergen, Norway, **13** Department of Pathology, Haukeland University Hospital, Bergen, Norway

Abstract

Context: Genetic testing for monogenic diabetes is important for patient care. Given the extensive genetic and clinical heterogeneity of diabetes, exome sequencing might provide additional diagnostic potential when standard Sanger sequencing-based diagnostics is inconclusive.

Objective: The aim of the study was to examine the performance of exome sequencing for a molecular diagnosis of MODY in patients who have undergone conventional diagnostic sequencing of candidate genes with negative results.

Research Design and Methods: We performed exome enrichment followed by high-throughput sequencing in nine patients with suspected MODY. They were Sanger sequencing-negative for mutations in the *HNF1A*, *HNF4A*, *GCK*, *HNF1B* and *INS* genes. We excluded common, non-coding and synonymous gene variants, and performed in-depth analysis on filtered sequence variants in a pre-defined set of 111 genes implicated in glucose metabolism.

Results: On average, we obtained 45 X median coverage of the entire targeted exome and found 199 rare coding variants per individual. We identified 0–4 rare non-synonymous and nonsense variants per individual in our *a priori* list of 111 candidate genes. Three of the variants were considered pathogenic (in *ABCC8*, *HNF4A* and *PPARG*, respectively), thus exome sequencing led to a genetic diagnosis in at least three of the nine patients. Approximately 91% of known heterozygous SNPs in the target exomes were detected, but we also found low coverage in some key diabetes genes using our current exome sequencing approach. Novel variants in the genes *ARAP1*, *GLIS3*, *MADD*, *NOTCH2* and *WFS1* need further investigation to reveal their possible role in diabetes.

Conclusion: Our results demonstrate that exome sequencing can improve molecular diagnostics of MODY when used as a complement to Sanger sequencing. However, improvements will be needed, especially concerning coverage, before the full potential of exome sequencing can be realized.

Citation: Johansson S, Irgens H, Chudasama KK, Molnes J, Aerts J, et al. (2012) Exome Sequencing and Genetic Testing for MODY. PLoS ONE 7(5): e38050. doi:10.1371/journal.pone.0038050

Editor: Ludmila Prokunina-Olsson, National Cancer Institute, National Institutes of Health, United States of America

Received: September 6, 2011; **Accepted:** May 2, 2012; **Published:** May 25, 2012

Copyright: © 2012 Johansson et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Funding: The study was supported in part by funds from the Research Council of Norway, the University of Bergen, Haukeland University Hospital, Helse Vest, Innovest, and the United States Public Health Service. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: pal.njolstad@uib.no

Introduction

MODY (maturity-onset diabetes of the young) is a heterogeneous group of diabetes caused by single gene defects in at least ten genes affecting pancreas development and beta-cell function [1,2,3]. The most common MODY forms are caused by mutations in the glucokinase gene (*GCK*) [4] and the hepatocyte transcription factor genes *HNF1A* and *HNF4A* [5,6]. *GCK*-MODY (MODY2) is a mild disease manifesting as slightly elevated fasting glucose, well controlled without medical treatment, and no risk for late diabetes-associated complications [7,8]. In contrast, *HNF1A*- and *HNF4A*-

MODY (MODY3 and MODY1, respectively) typically lead to progressive beta-cell dysfunction and high risk for late complications and patients often benefit from sulfonylurea treatment [9,10,11]. *HNF1B*-mutations result in a syndromic diabetes form (MODY5), which includes renal failure, genital and pancreatic malformations, and liver dysfunction [12,13]. According to the OMIM database, mutations in seven other genes (*BLK*, *CEL*, *INS*, *KLF11*, *NEUROD1*, *PAX4*, *PDX1*) can cause inherited diabetes with a MODY phenotype. There are also other forms of monogenic diabetes such as neonatal diabetes that presents before

six months of age and syndromic diabetes, in which other features than diabetes dominates the clinical pictures (reviewed in [1]).

Genetic testing in monogenic diabetes is important for diagnosis and treatment [1,2,3]. When MODY is suspected, the current approach involves PCR amplification and Sanger sequencing of candidate genes, frequently with an iterative approach based on clinical features. For example, most laboratories will first screen *HNFI1A*, followed by *HNFI4A* and *GCK* in subjects exhibiting the classical features of MODY; and first *GCK*, then *HNFI1A* and *HNFI4A*, if the diabetic phenotype is mild and fasting glucose 5.5–8.5 mmol/l [1,2]. If the patient presents with renal dysfunction, urogenital or pancreatic malformations, *HNFI1B* is usually the first gene that is tested [1,2].

Although systematic studies are lacking, our experience is that molecular genetic testing reveals a mutation in one of the common MODY genes in about 50% of probands referred to our laboratory. The remaining cases would also benefit from a genetic diagnosis, but the cost of sequencing other candidate genes often precludes further testing. A standard, complete investigation of *HNFI1A*, *HNFI4A* and *GCK* includes sequencing of 31 exons, where each sequencing reaction must be evaluated separately. Hence, the current approach is expensive and time-consuming, and establishes a molecular diagnosis only among a limited number of genes. Whole-exome capture and high-throughput sequencing has a great potential to detect causal gene variants in dominant and recessive disorders as well as in diseases due to *de novo* mutations [14,15,16,17,18]. Here, we describe our experience using exome sequencing in MODY patients referred to us for genetic testing.

Materials and Methods

Ethics Statement

The study was approved by the Regional Ethical Committee for Medical Research West and performed according to the Helsinki Declaration. We obtained verbal and written informed consent from the study participants.

Study Population

We carried out whole-exome sequencing on nine probands with MODY of unknown cause recruited from the Norwegian MODY Registry (Figure 1 and Table 1). There was diabetes running in the families for at least three generations, autosomal dominant inheritance and age at diagnosis 11–28 years for at least one family member. All probands were negative for mutations in *HNFI1A*, *HNFI4A*, *GCK*, *HNFI1B* and *INS* by Sanger sequencing.

Targeted Capture and Massive Parallel Sequencing

Targeted capture and massive parallel sequencing were performed at HudsonAlpha Institute for Biotechnology (Huntsville, AL) (File S1). In brief, SureSelect Human All Exon Kit (Agilent Technologies, Santa Clara, CA) was used for exome enrichment, and sequencing was performed on Genome Analyzer GAIIX (Illumina Inc., San Diego, CA). Samples were sequenced on one lane for paired-end 72-bp reads, and the samples with lowest yield were complemented with one additional lane of single-read 76-bp reads (P01, P03, P04, P05 P07).

Read Mapping and Variant Analysis

We mapped paired-end-reads and single-reads to the reference human genomes (UCSC NCBI37/hg19) using Burrows-Wheeler Alignment tool (BWA) [19] (see also supporting Materials and Methods File S1). PCR duplicates were removed with PICARD (<http://picard.sourceforge.net>) followed by base quality recalibra-

tion using GATK [20]. SNPs and indels were called by SAMtools [21] mpileup, correcting for overestimated mapping quality from BWA. SNPs were filtered by the following criteria: (1) SNPs should not be in a cluster with window-size of 10 bp, (2) sequencing depth should be at least 8 X; and (3) quality score should be ≥ 30 . We used Annovar [22] (Nov 22, 2011) and in-house scripts to annotate and filter variants after variant calling. The performance of our exome sequencing variant calling pipeline was tested against heterozygous genotypes present in the enrichment target regions derived from the Affymetrix 6.0 whole-genome genotyping array in seven of the individuals (File S1).

In silico Analyses of Candidate Variants

We evaluated possible functional significance of the variants using PolyPhen v2.0.23 [23], Align-GVGD (<http://agvgd.iarc.fr>) and SIFT (<http://sift.jcvi.org/>), see File S1.

Selection of genes of Interest

We selected genes previously implicated in monogenic diabetes and related syndromes [1,24], genes with important roles in the beta cell [25,26] and genes implicated from whole-genome-significant SNP associations with type 2 diabetes (T2D) or fasting glucose [24,27,28,29,30,31]. For the T2D and fasting glucose-associated regions, we selected named gene/genes highlighted in the respective publications. Notably, experimental evidence directly supporting that these genes are responsible for the associations, is mostly lacking. The 111 genes, of which 109 were included in the capture assay, totaled 272 kb of exonic sequence (Table S1).

Rare Variant Validation

All variants in the candidate mutation set were validated by PCR amplification of the variant-containing exon from the original patient sample, followed by Sanger sequencing. Frequency estimates were generated by genotyping 340 Norwegian healthy controls using the MassARRAY iPLEX system or by Sanger sequencing.

Results

By sequencing the exomes of the nine MODY probands (Table 1, Figure 1), we obtained 3.5–5.8 Gb mapable sequence per sample with 36–57 X median coverage of the targeted exome and 88–93% of the exome targeted at least eight times (Table 2). The candidate diabetes genes showed similar coverage (Tables S1 and S2). The exceptions were *DGKB* and *THADA*, which were not present on the exome enrichment array highlighting one problem with our approach: current target capture reagents may not include all exons of interest. The MODY genes *HNFI4A* and *HNFI1B* showed relatively good coverage ($\geq 87\%$, at 8X) throughout the entire coding regions, while *GCK* (83%), *HNFI1A* (72%) and *INS* (58%) were less uniformly covered (Tables S1 and S2).

We identified an average of 14,463 substitutions and indels per sample (in the targeted exome) after quality control (Table 3). The quality of the data was investigated by comparisons from 7,800 genotyped SNPs, present in the Agilent capture region, and obtained from the Affymetrix 6.0 genotyping array for seven of the nine individuals. Between 89 and 92% of the heterozygous genotyping-array-SNPs present in the capture regions were detected (File S1 and Table S3).

Next, we developed a data reduction pipeline consisting of several steps (Table 3). We first excluded all variants not present in the actual coding sequence or in splice sites; and synonymous variants other than those occurring at canonical splice sites. We

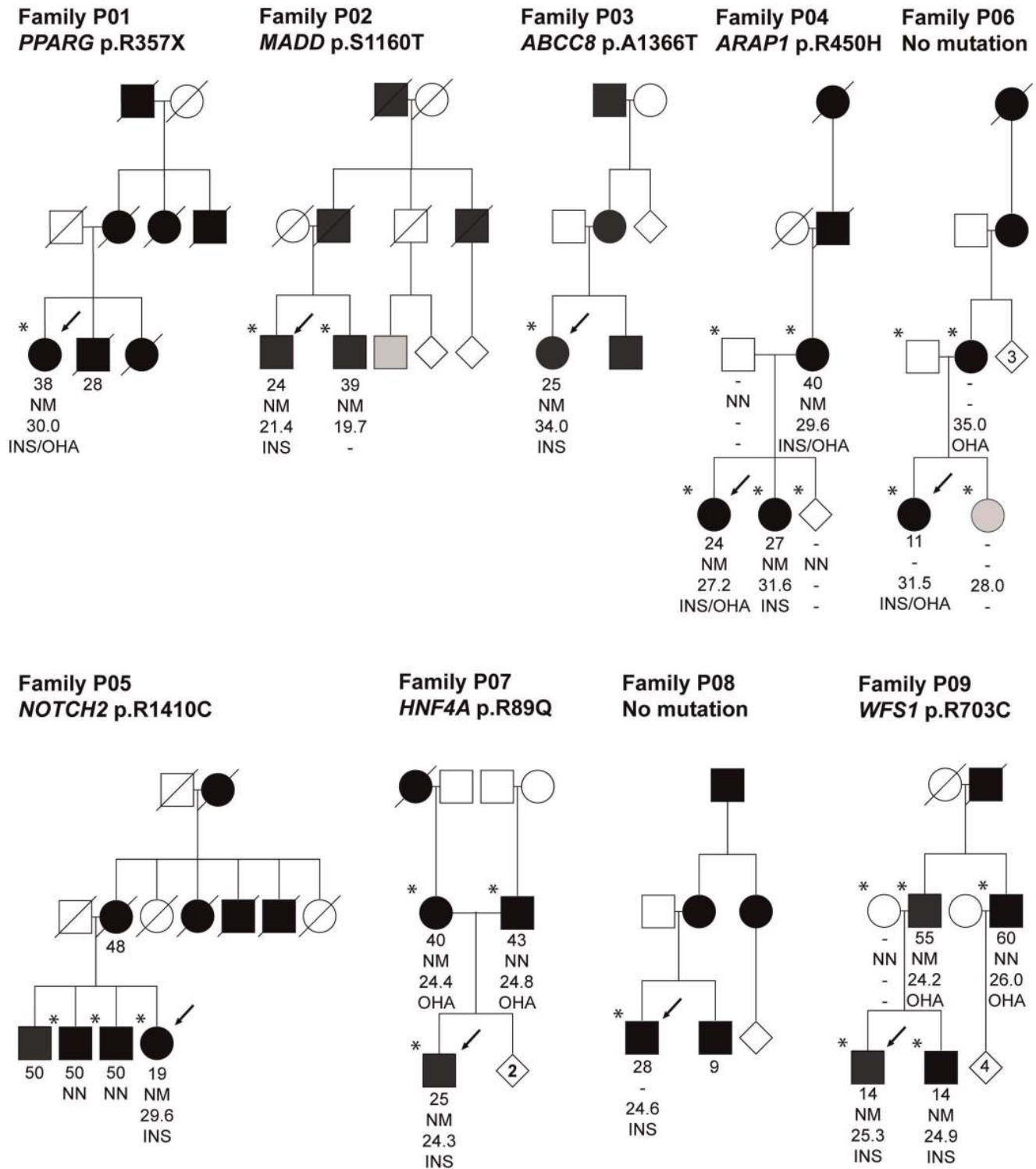


Figure 1. Partial pedigrees of the nine MODY families investigated. Squares represent male family members, circles female, and diamonds sex unknown. Numbers inside diamonds show the number of siblings. Black and grey symbols represent persons with diabetes and impaired glucose tolerance, respectively. An arrow denotes the proband in each family and stars indicate those subjects for whom DNA was available. Data under the symbols represent from top to bottom: age at diabetes diagnosis, mutation carrier status (N = Normal allele, M = Mutation), BMI and treatment (INS = Insulin, OHA = Oral Hypoglycemic Agents). doi:10.1371/journal.pone.0038050.g001

subsequently filtered against an in-house database of genetic variants from 50 Norwegian whole exomes and finally excluded variants with minor allele frequencies >0.5% of the 1000

Genomes Project. This reduced the number to 183–213 rare, coding single-nucleotide substitutions and coding indels per individual (Table 3). Combining all nine individuals, this

Table 1. Clinical characteristics of the nine MODY probands investigated.

Proband	P01 ^a	P02	P03	P04	P05	P06	P07	P08	P09
Sex	F	M	F	F	F	F	M	M	M
Age of diabetes diagnosis (y)	38 ^b	24	25	24	19	11	25	28	14
Number of generations with DM	3	3	3	4	3	4	3	3	3
Current status									
Age (y)	68	52	38	31	55	16	35	35	29
BMI	30.1	21.4	29.9	27.2	28.3	31.5	24.3	24.6	25.3
Insulin dose (U/kg/day)	2.2	0.4	0	0.8	0.7	2.0	0.8	NA	0.5
OHA	Yes	No	Yes	Yes	No	Yes	No	No	No
Glycosylated hemoglobin (%)	7.5	6.1	9.0	6.7	7.1	8.4	7.2	9.3	6.4
Other clinical features									
Hypertension	Yes	No	No	No	Yes	Yes	Yes	No	No
Hypercholesterolemia	Yes	No	No	No	Yes	Yes	No	No	No
Retinopathy	Yes	No	No	No	Yes	No	Yes	No	No
Nephropathy	Yes	No	No	No	No	No	No	No	No
Arteriosclerosis	Yes	No	No	No	No	No	No	No	No
Polyneuropathy	Yes	No	No	No	Yes	No	No	No	No

None of the individuals had hearing loss or acanthosis nigricans. Data regarding hepatic steatosis, renal cysts, polycystic ovarian disease and triglyceride status were not available in all individuals.

^aP01 also had partial lipodystrophy with reduced subcutaneous fat in extremities and excess of abdominal subcutaneous fat. There was no evidence of hypertriglyceridemia, hepatic steatosis, acanthosis nigricans or polycystic ovarian disease.

^bAn affected family member was diagnosed at age 28 years.

Abbreviations: BMI, body mass index; DM, diabetes mellitus; NA, not available; OHA, Oral Hypoglycemic agents; U, international unit (used for insulin doses); y, year; kg, kilo gram body weight.

doi:10.1371/journal.pone.0038050.t001

resulted in 1,733 different variants located in 1,569 different genes. Only 50 variants were present in more than one of the nine individuals. On the gene level, 266 genes were listed with rare (most often different) variants in more than one individual, and 24 genes with rare variants in three or more individuals. Hence, despite our rigid procedure for variant filtration, a large number of potential candidate genes emerged from the nine patient data sets.

We then focused on the candidate gene list (Table S1). We identified 14 rare coding variants in 12 genes of the 111

candidate genes. Thirteen of the variants were verified by Sanger sequencing (Tables 3 and 4). Frequency estimates in 340 healthy controls, computational methods to estimate deleteriousness and literature searches were performed for the remaining 13 variants (Table 4). When available, additional family members were sequenced for variants not present in either the 1000 G database, dbSNP or in our 340 Norwegian controls (Figure 1).

Table 2. Overall exome coverage and target gene set coverage statistics.

Sample ID	Unique Gbs aligned	Median (X)	Percentage at ≥		Percentage at ≥		
			8X	20X	Median (X)	8X	20X
P01	3.5	36	89	71	33	85	66
P02	3.5	36	88	71	34	86	68
P03	4.8	50	92	79	46	89	75
P04	5.5	57	92	81	52	90	77
P05	5.4	55	92	8	49	89	76
P06	3.6	39	89	73	35	85	68
P07	5.8	59	93	82	55	90	78
P08	3.7	39	89	73	37	86	69
P09	3.5	37	89	71	35	86	68
Average	4.4	45	90	76	42	87	72

doi:10.1371/journal.pone.0038050.t002

Table 3. Overview of all substitutions and indels detected in the nine probands before and after variant reduction.

Variants and filter	Patients									Average
	P01	P02	P03	P04	P05	P06	P07	P08	P09	
All exonic	14102	14219	14773	14974	14668	14253	14720	14190	14269	14463
Exonic coding	6460	6582	6776	6855	6634	6526	6744	6557	6565	6633
Not in in-house database (50 samples)	250	215	263	255	252	255	264	267	260	253
Not in 1000 G >0.5%	193	183	197	202	197	197	202	213	206	199
Candidate variants in 111 target genes	1	2	3	1	2	0	1	0	4	2

doi:10.1371/journal.pone.0038050.t003

Exome Sequencing Reveals Three Variants in genes Known to Cause Autosomal Dominant Disease

In subject P01, we identified the heterozygous nonsense mutation c.1071G>A/p.R357X introducing a premature stop codon in *PPARG* exon 7. This mutation has previously been shown to cause severe insulin-resistant diabetes and partial lipodystrophy [32,33,34]. The proband's age at diagnosis was 38 years; she was included in our study because a family member was diagnosed at age 28 years. The proband's BMI was 27.6 kg/m² in 2002, 30.1 kg/m² in 2010, and her insulin requirement has in the same period increased from 1.4 to 2.2 U/kg/day. There was a high prevalence of micro- and macro-vascular complications in most of the affected family members, although none were available for genetic testing (Table 1, Figure 1). Thus, at recruitment the proband had a MODY phenotype, but was insulin-resistant on follow-up. The same mutation has been reported in patients with

a similar clinical picture [32,33,34]. We consider this mutation pathogenic.

In subject P03, we detected the novel and heterozygous non-synonymous *ABCC8* mutation c.4096G>A/p.A1366T. Amino acid 1366 is highly conserved and located in the ATP-binding domain. Other nearby amino acid substitutions are associated with either congenital hyperinsulinism or neonatal/adult-onset diabetes [35,36]. The proband was diagnosed with diabetes at 25 years of age and is currently treated with sulfonylurea and metformin. All four diabetic family members (none available for genetic testing) were treated with OHA. The proband's age at diagnosis was late for *ABCC8* diabetes although other such cases have been described [35,37,38,39]. The large size of *ABCC8* makes it less amenable to Sanger-based mutation screening, which may underestimate the role of this gene in MODY. Since the proband was sulfonylurea-sensitive, we categorized p.A1366T as being probably pathogenic. Further sequencing studies of unselected MODY patients will

Table 4. Rare coding variants identified in the 111 target candidate genes using whole exome sequencing in nine patients with suspected MODY.

Gene	Chr:Position	Variant	dbSNP132/1000 G ^b frequency	Frequency in 340 Norwegian controls	SIFT/ PolyPhen-2/ AlignGVGD ^a	Patient	Conclusion
<i>ABCC8</i>	11:17418486	c.4096G>A/p.A1366T	-/0	0	-+/C55	P03	Pathogenic
<i>ALMS1</i>	2:73677199	c.3542C>T/p.T1181I	-/0	0.1%	n.a./n.a.	P09	
<i>ARAP1</i>	11:72421497	c.1349G>A/p.R450H	-/0	0	+/-/C0	P04	
<i>CRY2</i>	11:45893711	c.1528G>C/p.G510R	-/0	0.1%	-/-/C15	P02	
<i>GLIS3</i>	9:4286332	c.94C>G/p.R32G	-/0	0	+/-/C0	P03	
<i>HADH</i>	4:108940732	c.456G>T/p.Q152H	rs1051519/0.2%	-	-/-/C0	P09	
<i>HNF4A</i>	20:43034848	c.266G>A/p.R89Q	-/0	0	+/-/C35	P07	Pathogenic
<i>MADD</i>	11:47317569	c.3479G>C/p.S1160T	-/0	0	-/+/-/C55	P02	
<i>NOTCH2</i>	1:120468211	c.4228C>T/p.R1410C	-/0	0	+/-/C25	P05	
	1:120478125	c.3625T>G/p.F1209V	-/0	0.4%	+/-/C45	P03	
	1:120548095	c.272G>T/p.R91L	FALSE	FALSE	FALSE	P05	False positive
<i>PPARG</i>	3:1258536	c.1071G>A/p.R357X	-/0	0	Nonsense	P01	Pathogenic
<i>SREBF1</i>	17:17718592	c.2435G>A/p.R812Q	-/0	1.0%	-/-/C0	P09	
<i>WFS1</i>	4:6354530	c.2107C>T/p.R703C	-/0	0	+/-/C65	P09	

^aSIFT: - tolerated, + not tolerated/PolyPhen-2: - benign, + possibly damaging, ++ probably damaging/Align-GVGD: the Grantham variation (GV), and the Grantham deviation (GD) are combined to provide graded classifiers from most likely to interfere with function (class C65) to least likely (class C0).

^bAllele frequencies from the interim analysis of phase I of the 1000 Genomes Project, 2010.08.04 sequence index, which included 629 samples (SNPs released in November 2010, indels released in February 2011).

Abbreviation: Chr, chromosome number; 1000 G, the 1000 Genomes Project; n.a., not analysed due to insufficient number of alignments to make prediction.

doi:10.1371/journal.pone.0038050.t004

elucidate if *ABCC8* is a more common cause of MODY than previously anticipated.

In subject P07, we identified a novel non-synonymous mutation c.266G>A/p.R89Q in *HNF4A*. This was surprising since *HNF4A* already had been screened. Sanger re-sequencing confirmed the mutation. When re-examining the first electropherogram, the mutation was detectable. Hence, it had been overlooked. Another substitution of the same codon, c.265C>T/p.R89W, has been identified in MODY [40]. The amino acid residue at position 89 of *HNF4A* is highly conserved from *Drosophila* to humans and part of the DNA binding domain. Both parents are of normal weight but developed diabetes in their early forties (Figure 1). Sanger sequencing revealed that the mutation was inherited from the maternal side of the family that appears to have a stronger history of diabetes. After receiving the molecular diagnosis the proband made a successful transfer from insulin to sulfonylurea. We consider the mutation p.R89Q pathogenic.

Other Rare Variants in the Candidate Gene Set

In the 111 candidate genes, we also identified novel (not present in 50 in-house exomes, 340 healthy controls or 1000 Genomes) variants in potentially interesting genes implicated in susceptibility to diabetes, albeit thus far not in an autosomal dominant mode of inheritance: *ARAPI*, *GLIS3*, *MADD*, *NOTCH2* and *WFS1* (Table 4). Each genetic variant is discussed in some detail below.

In subject P02, we detected the novel and heterozygous non-synonymous *MADD* variant c.3479G>C/p.S1160T. The *MADD* (MAP-kinase activating death domain) protein is known to have a role in apoptosis [41] and SNPs in the *MADD* region are associated with elevated pro-insulin and fasting glucose levels [42]. The subject was diagnosed with diabetes at 24 years of age. He has for ten years been treated with sulfonylurea and is currently on insulin (0.5 U/kg/day). C-peptide and proinsulin were detectable, however, not elevated. The affected and lean brother also carried the variant. He was diagnosed with diabetes 39 years old and is treated with oral hypoglycaemic agents (OHA). No other family members were available for genetic analysis.

Individual P03 (who had a probably pathogenic *ABCC8* mutation) also had a potentially interesting variant in *GLIS3* which is a transcription factor expressed in beta-cells and important for insulin gene expression [43]. Mutations in *GLIS3* can cause a recessive form of neonatal diabetes and congenital hypothyroidism (OMIM#610199) [44,45]. There are, however, no reports on dominant *GLIS3* mutations, and the variant is located in a protein region with no known function. With only the proband available for genetic testing, it was not possible to study the segregation of the variant in the family. We consider that the *ABCC8* mutation is more likely to be the pathogenic variant in this patient.

In subject P04, we detected the *ARAPI* (previously *CENTD2*) variant c.1349G>A/p.R450H. It is predicted to be benign by PolyPhen and AlignGVGD (Table 4) but damaging by SIFT, and it co-segregates with diabetes in the core family (Figure 1). Common variants at this locus have been associated with type 2 diabetes, fasting glucose and pro-insulin level [28,46,47], and it is suggested that this effect is mediated through reduced insulin secretion capacity [46,47]. However, as with most GWAS-associated regions, the causative variant has not yet been pinpointed, and the nearby *STARD10* gene was recently suggested as a better biological candidate gene in the region [47]. Although not on our original candidate gene list, in retrospect, no rare variants were detected in *STARD10*.

The *NOTCH2* variant c.4228C>T/p.R1410C, found in P05, did not co-segregate with diabetes (Figure 1). This gene was

implicated as a type 2 diabetes locus in a recent GWAS meta-analysis [28]. It is also known that heterozygous mutations in *NOTCH2* can cause Alagille syndrome (OMIM#610205) and Hajdu-Cheney syndrome (OMIM #102500). The patient showed no symptoms suggesting any of these diseases.

In subject P09, we detected the novel *WFS1* variant c.2107C/T/p.R703C. The affected amino acid residue is strongly conserved and the variant is suggested to be probably damaging by all three prediction programs (Table 4). Recessive mutations in *WFS1* can lead to Wolfram syndrome (OMIM #222300), which includes diabetes, hearing impairment and psychiatric disease, while heterozygous carriers appear to show no major symptoms associated with diabetes. *Wfs1* null mice and genetic association studies suggest a role in insulin secretion [48,49]. The proband developed diabetes 14 years old with no type 1 auto-antibodies and currently requires 0.5 U/kg/day insulin. There was no familial hearing impairment. The affected brother and father carried the variant, but not the affected uncle (Figure 1). Age-of-diagnosis and insulin requirements are distinctly different between the affected brothers and their father and uncle.

Discussion

Exome sequencing has shown a great potential for identification of disease mutations in monogenic disorders [14,15,16,17,18]. However, it is not clear how representative the early proof-of-principle studies are and whether this technology is ready to replace or complement traditional Sanger sequencing for clinical genetic testing.

Here, we show that exome sequencing can provide a significant diagnostic advantage in a substantial fraction of patients where Sanger sequencing often is inefficient, such as cases with atypical clinical presentation (family P03) or when clinical information is limited (family P01). For this group of patients, exome sequencing is an attractive option compared to the current “phenotypically” driven genetic testing as it allows testing beyond the short list of genes typically tested by Sanger sequencing.

To illustrate how this technology could be utilized for routine diagnostic use, we restricted our analysis to a list of 111 candidate genes. Our list included known disease genes for monogenic diabetes, insulin resistance and diseases related to glucose homeostasis. We also explored rare variants in candidates such as genes encoding transcription factors important for pancreatic development and islet specification/differentiation [25,26], and genes identified in GWAS of diabetes and fasting glucose levels. Our study identified some rare variants in the latter gene categories (Table 4 and results section). Although these variants are located in attractive candidates, to claim causality would obviously require much more extensive proof than for genes already known to cause autosomal-dominant diabetes. Such evidence would include co-segregation between variant and disease in large families, the presence in other subjects with a similar phenotype together with functional and clinical studies. Especially the limited number of available family members for segregation analysis, does not allow us to reach this level of support for the variants in our “candidate gene” category. Hence, we could not determine whether these variants are causing MODY or at least may act as polygenic risk factors that warrant further investigations (Table 4).

A possible advantage with exome sequencing is that it also allows for an extensive search for completely novel diabetes genes in individuals with no genetic defect in the known diabetes genes. However, for diabetes, where the genes for several monogenic forms already have been detected [1], the search for remaining,

unmapped disease loci will be hampered by significant locus- and clinical heterogeneity. Our registry-based clinical sample with limited access to extended pedigrees was not powered to identify novel disease genes among the approximately 200 rare coding variants in each individual. Thus, international efforts to sequence the entire exomes of larger numbers of carefully selected subjects and to identify large multi-generational diabetes families may be a way forward.

For diagnostic utility, our study reveals that exome sequencing can increase the possibility for a genetic diagnosis in MODY. The coverage for certain key genes must, however, be improved before exome sequencing can replace Sanger sequencing in routine molecular diagnostics. Recent and ongoing improvements in capture hybridization and high-throughput sequencing technologies are promising, but the coverage problem may not be solved completely by new enrichment kits, higher read depths and longer reads. In the meantime, it might be attractive to use tailored hybridization capture for the disease of interest followed by very high-coverage sequencing of the disease-specific gene panels [50]. The increased coverage for the target genes must be weighted against the cost of developing, optimizing and keeping up-to-date disease-specific gene panels and the limited ability to detect unexpected phenotype-genotype correlations.

In conclusion, we consider phenotypically driven Sanger sequencing still as the first choice for genetic testing in patients with classical features of MODY. Exome sequencing is currently an important complement when Sanger sequencing is negative, or in patients with atypical clinical presentation. In the near future, we believe that tailored hybridization capture for selected genes of interest and very high-coverage sequencing of specific gene panels will replace Sanger sequencing. Ongoing refinements in the design of capture reagents, sequencing technologies and bioinformatics will, however, most likely ultimately lead to exome and possibly whole-genome sequencing as state-of-the art in molecular diagnostics of MODY.

References

- Molven A, Njolstad PR (2011) Role of molecular genetics in transforming diagnosis of diabetes mellitus. *Expert Rev Mol Diagn* 11: 313–320.
- Hattersley A, Bruining J, Shield J, Njolstad P, Donaghue KC (2009) The diagnosis and management of monogenic diabetes in children and adolescents. *Pediatr Diabetes* 10 Suppl 12: 33–42.
- Murphy R, Ellard S, Hattersley AT (2008) Clinical implications of a molecular genetic classification of monogenic beta-cell diabetes. *Nat Clin Pract Endocrinol Metab* 4: 200–213.
- Vionnet N, Stoffel M, Takeda J, Yasuda K, Bell GI, et al. (1992) Nonsense mutation in the glucokinase gene causes early-onset non-insulin-dependent diabetes mellitus. *Nature* 356: 721–722.
- Yamagata K, Furuta H, Oda N, Kaisaki PJ, Menzel S, et al. (1996) Mutations in the hepatocyte nuclear factor-4[alpha] gene in maturity-onset diabetes of the young (MODY1). *Nature* 384: 458–460.
- Yamagata K, Oda N, Kaisaki PJ, Menzel S, Furuta H, et al. (1996) Mutations in the hepatocyte nuclear factor-1[alpha] gene in maturity-onset diabetes of the young (MODY3). *Nature* 384: 455–458.
- Froguel P, Zouali H, Vionnet N, Velho G, Vaxillaire M, et al. (1993) Familial Hyperglycemia Due to Mutations in Glucokinase – Definition of a Subtype of Diabetes Mellitus. *New England Journal of Medicine* 328: 697–702.
- Osbak KK, Colclough K, Saint-Martin C, Beer NL, Bellanné-Chantelot C, et al. (2009) Update on mutations in glucokinase (GCK), which cause maturity-onset diabetes of the young, permanent neonatal diabetes, and hyperinsulinemic hypoglycemia. *Human Mutation* 30: 1512–1526.
- Sovik O, Njolstad P, Følling I, Sagen J, Cockburn BN, et al. (1998) Hyperexcitability to sulphonylurea in MODY3. *Diabetologia* 41: 607–608.
- Pearson ER, Pruhova S, Tack CJ, Johansen A, Castleden HAJ, et al. (2005) Molecular genetics and phenotypic characteristics of MODY caused by hepatocyte nuclear factor 4 α mutations in a large European collection. *Diabetologia* 48: 878–885.
- Steele AM, Shields BM, Shepherd M, Ellard S, Hattersley AT, et al. (2010) Increased all-cause and cardiovascular mortality in monogenic diabetes as a result of mutations in the HNF1A gene. *Diabetic Medicine* 27: 157–161.

Supporting Information

File S1 Supplementary Materials and Methods.
(DOC)

Table S1 Candidate genes with reason for inclusion and average coverage in the nine tested samples.
(DOC)

Table S2 Fraction of target bases covered at minimum 8 X for each sample and gene.
(DOC)

Table S3 Comparison between heterozygous genotypes obtained from the Affymetrix 6.0 chip and exome sequencing.
(DOC)

Acknowledgments

We thank the patients and their families for participation in the study, Helge Ræder and Jørn V. Sagen for help regarding the Norwegian MODY Registry, Braden Boone (Hudson Alpha) for generation of the raw exome sequencing data; Monika Ringdal, Benedikte Rosenlund and Louise Greve for Sanger sequencing and mutation validation, Bjørn-Ivar Haukanes for discussions on *in silico* mutational characterization, Matthew Hurles for hosting Stefan Johansson during his research stay at the Wellcome Trust Sanger Institute and Carol Scott (Wellcome Trust Sanger Institute) for sharing her expertise and scripts related to data management.

Author Contributions

Conceived and designed the experiments: SJ HI KKC JM PK PRN JA FSR IJ SL KL GIB AM. Performed the experiments: SJ HI KKC JM PK PRN JA FSR SL KL. Analyzed the data: SJ HI KKC JM JA FSR. Contributed reagents/materials/analysis tools: SJ HI KKC JM PK PRN JA FSR IJ SL. Wrote the paper: SJ HI KKC JM PK PRN JA FSR IJ SL KL GIB AM.

- Horikawa Y, Iwasaki N, Hara M, Furuta H, Hinokio Y, et al. (1997) Mutation in hepatocyte nuclear factor-1 beta gene (TCF2) associated with MODY. *Nat Genet* 17: 384–385.
- Lindner TH, Njolstad PR, Horikawa Y, Bostad L, Bell GI, et al. (1999) A novel syndrome of diabetes mellitus, renal dysfunction and genital malformation associated with a partial deletion of the pseudo-POU domain of hepatocyte nuclear factor-1beta. *Hum Mol Genet* 8: 2001–2008.
- Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, et al. (2010) Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. *Nat Genet* 42: 790–793.
- Ng SB, Buckingham KJ, Lee C, Bigham AW, Tabor HK, et al. (2010) Exome sequencing identifies the cause of a mendelian disorder. *Nat Genet* 42: 30–35.
- Vissers LELM, de Ligt J, Gilissen C, Janssen I, Stehouwer M, et al. (2010) A de novo paradigm for mental retardation. *Nat Genet* 42: 1109–1112.
- Bolze A, Byun M, McDonald D, Morgan NV, Abhyankar A, et al. (2010) Whole-Exome-Sequencing-Based Discovery of Human FADD Deficiency. *American journal of human genetics* 87: 873–881.
- Bonnefond A, Durand E, Sand O, De Graeve F, Gallina S, et al. (2010) Molecular diagnosis of neonatal diabetes mellitus using next-generation sequencing of the whole exome. *PLoS One* 5: e13630.
- Durbin RM, Abecasis GR, Altshuler DL, Auton A, Brooks LD, et al. (2010) A map of human genome variation from population-scale sequencing. *Nature* 467: 1061–1073.
- McKenna A, Hanna M, Banks E, Sivachenko A, Cibulskis K, et al. (2010) The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res* 20: 1297–1303.
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, et al. (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25: 2078–2079.
- Wang K, Li M, Hakonarson H (2010) ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Research* 38: e164.
- Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, et al. (2010) A method and server for predicting damaging missense mutations. *Nat Methods* 7: 248–249.

24. McCarthy MI (2010) Genomics, type 2 diabetes, and obesity. *N Engl J Med* 363: 2339–2350.
25. Edghill EL, Minton JA, Groves CJ, Flanagan SE, Patch AM, et al. (2010) Sequencing of candidate genes selected by beta cell experts in monogenic diabetes of unknown aetiology. *JOP* 11: 14–17.
26. Oliver-Krasinski JM, Stoffers DA (2008) On the origin of the beta cell. *Genes Dev* 22: 1998–2021.
27. Voight BF, Scott LJ, Steinthorsdottir V, Morris AP, Dina C, et al. (2010) Twelve type 2 diabetes susceptibility loci identified through large-scale association analysis. *Nat Genet* 42: 579–589.
28. Zeggini E, Scott LJ, Saxena R, Voight BF, Marchini JL, et al. (2008) Meta-analysis of genome-wide association data and large-scale replication identifies additional susceptibility loci for type 2 diabetes. *Nat Genet* 40: 638–645.
29. Dupuis J, Langenberg C, Prokopenko I, Saxena R, Soranzo N, et al. (2010) New genetic loci implicated in fasting glucose homeostasis and their impact on type 2 diabetes risk. *Nat Genet* 42: 105–116.
30. Feero WG, Guttmacher AE, McCarthy MI (2010) Genomics, Type 2 Diabetes, and Obesity. *New England Journal of Medicine* 363: 2339–2350.
31. Grarup N, Sparso T, Hansen T (2010) Physiologic Characterization of Type 2 Diabetes-Related Loci. *Current Diabetes Reports* 10: 485–497.
32. Barroso I, Gurnell M, Crowley VEF, Agostini M, Schwabe JW, et al. (1999) Dominant negative mutations in human PPAR[γ] associated with severe insulin resistance, diabetes mellitus and hypertension. *Nature* 402: 880–883.
33. Agostini M, Schoenmakers E, Mitchell C, Szatmari I, Savage D, et al. (2006) Non-DNA binding, dominant-negative, human PPAR[γ] mutations cause lipodystrophic insulin resistance. *Cell Metabolism* 4: 303–311.
34. Gurnell M (2007) ‘Striking the Right Balance’ in Targeting PPAR[γ] in the Metabolic Syndrome: Novel Insights from Human Genetic Studies. *PPAR Res* 2007: 83593. 83593 p.
35. Babenko AP, Polak M, Cavé H, Busiah K, Czernichow P, et al. (2006) Activating Mutations in the ABCC8 Gene in Neonatal Diabetes Mellitus. *New England Journal of Medicine* 355: 456–466.
36. Flanagan SE, Clauin S, Bellanné-Chantelot C, de Lonlay P, Harries LW, et al. (2009) Update of mutations in the genes encoding the pancreatic beta-cell KATP channel subunits Kir6.2 (KCNJ11) and sulfonylurea receptor 1 (ABCC8) in diabetes mellitus and hyperinsulinism. *Human Mutation* 30: 170–180.
37. Porksen S, Laborie L, Nielsen L, Louise Max Andersen M, Sandal T, et al. (2010) Disease progression and search for monogenic diabetes among children with new onset type 1 diabetes negative for ICA, GAD- and IA-2 Antibodies. *BMC Endocrine Disorders* 10: 16.
38. Bowman P, Flanagan S, Edghill E, Damhuis A, Shepherd M, et al. (2012) Heterozygous ABCC8 mutations are a cause of MODY. *Diabetologia* 55: 123–127.
39. Riveline JP, Rousseau E, Reznik Y, Fetita S, Philippe J, et al. (2012) Clinical and metabolic features of adult-onset diabetes caused by ABCC8 mutations. *Diabetes Care* 35: 248–251.
40. Harries LW, Locke JM, Shields B, Hanley NA, Hanley KP, et al. (2008) The diabetic phenotype in HNF4A mutation carriers is moderated by the expression of HNF4A isoforms from the P1 promoter during fetal development. *Diabetes* 57: 1745–1752.
41. Efimova EV, Al-Zoubi AM, Martinez O, Kaithamana S, Lu S, et al. (2004) IG20, in contrast to DENN-SV, (MADD splice variants) suppresses tumor cell survival, and enhances their susceptibility to apoptosis and cancer drugs. *Oncogene* 23: 1076–1087.
42. Ingelsson E, Langenberg C, Hivert MF, Prokopenko I, Lyssenko V, et al. (2010) Detailed physiologic characterization reveals diverse mechanisms for novel genetic Loci regulating glucose and insulin metabolism in humans. *Diabetes* 59: 1266–1275.
43. Yang Y, Chang BH-J, Samson SL, Li MV, Chan L (2009) The Krüppel-like zinc finger protein Glis3 directly and indirectly activates insulin gene transcription. *Nucleic Acids Research* 37: 2529–2538.
44. Senee V, Chelala C, Duchatelet S, Feng D, Blanc H, et al. (2006) Mutations in GLIS3 are responsible for a rare syndrome with neonatal diabetes mellitus and congenital hypothyroidism. *Nat Genet* 38: 682–687.
45. Dimitri P, Warner J, Minton J, Patch A-M, Ellard S, et al. (2010) Novel GLIS3 mutations demonstrate an extended multisystem phenotype. *Eur J Endocrinol*. pp EJE-10-0893.
46. Nielsen T, Sparso T, Grarup N, Jorgensen T, Pisinger C, et al. (2011) Type 2 diabetes risk allele near CENTD2 is associated with decreased glucose-stimulated insulin release. *Diabetologia* 54: 1052–1056.
47. Strawbridge RJ, Dupuis Je, Prokopenko I, Barker A, Ahlqvist E, et al. (2011) Genome-Wide Association Identifies Nine Common Variants Associated With Fasting Proinsulin Levels and Provides New Insights Into the Pathophysiology of Type 2 Diabetes. *Diabetes* 60: 2624–2634.
48. Sandhu MS, Weedon MN, Fawcett KA, Wasson J, Debenham SL, et al. (2007) Common variants in WFS1 confer risk of type 2 diabetes. *Nat Genet* 39: 951–953.
49. Cheurfa N, Brenner G, Reis A, Dubois-Laforgue D, Roussel R, et al. (2010) Decreased insulin secretion and increased risk of type 2 diabetes associated with allelic variations of the WFS1 gene: the Data from Epidemiological Study on the Insulin Resistance Syndrome (DESIR) prospective study. *Diabetologia*. pp 1–9.
50. Audo I, Bujakowska K, Leveillard T, Mohand-Said S, Lancelot M-E, et al. (2012) Development and application of a next-generation-sequencing (NGS) approach to detect known and novel gene defects underlying retinal diseases. *Orphanet Journal of Rare Diseases* 7: 8.