Optimal management of diabetes is essential to avoid chronic complications. Maturity-onset diabetes of the young (MODY) is the most common type of monogenic diabetes. It constitutes a group of early-onset autosomal dominant forms of diabetes that together account for up to 2% of all diabetes cases.

INTRODUCTION

Monogenic diabetes constitutes a heterogeneous group of single-gene disorders. Its diagnosis helps understand the pathogenesis of the disease, defines the risk of diabetes within a family, and enables to modify treatment in a substantial number of patients. Genetic testing for monogenic diabetes using targeted next-generation sequencing in patients with maturity-onset diabetes of the young

Magdalena Szopa1,2*, Agnieszka Ludwig-Gałęzowska1,3*, Piotr Radkowski3, Jan Skupień1,3, Barbara Zapała4, Teresa Płatek4, Tomasz Klupa1,2, Beata Kiec-Wilk1,2, Maciej Borowiec5,6, Wojciech Młynarski6, Paweł Wołkowski2, Maciej T. Małecki1,2

1 Department of Metabolic Diseases, Jagiellonian University Medical College, Kraków, Poland
2 Department of Metabolic Diseases, University Hospital, Kraków, Poland
3 Center for Medical Genomics OMICRON, Jagiellonian University Medical College, Kraków, Poland
4 Department of Clinical Biochemistry, Jagiellonian University Medical College, Kraków, Poland
5 Department of Clinical and Laboratory Genetics, Medical University of Lodz, Łódź, Poland
6 Department of Pediatrics, Oncology, Hematology, and Diabetology, 1st Department of Pediatrics, Medical University of Lodz, Łódź, Poland

ABSTRACT

INTRODUCTION Molecular diagnosis of monogenic diabetes mellitus is important for individualized patient care. Next-generation sequencing (NGS) enables a simultaneous analysis of multiple genes in a single test. OBJECTIVES We aimed to assess the feasibility of using NGS for detecting mutations in a set of known monogenic diabetes gene mutations in a cohort of Polish patients with maturity-onset diabetes of the young (MODY) with earlier negative Sanger sequencing results for HNF1A-MODY or GCK-MODY. PATIENTS AND METHODS We selected a panel of 28 chromosomal genes in which mutations have been reported to cause monogenic diabetes. The MiSeq platform was used for NGS. An exon-capture assay was designed to include coding regions and splice sites. A total of 54 patients with existing negative Sanger sequencing screening results for HNF1A or GCK gene mutations were selected for the study. RESULTS NGS results were generated for all 54 patients and 9 positive controls with previously identified HNF1A or GCK gene mutation. All selected positive controls were confirmed by NGS. Among 28 genes, mutations were detected in 16. The type of the analyzed genetic changes was described in the NGS study as high (n = 3) or moderate (n = 76). Among the detected mutations, there were 4 known GCK gene mutations that had been previously missed in Sanger sequencing. So far, Sanger sequencing allowed us to confirm 21 gene mutations detected by NGS, and segregation with diabetes in 14 pedigrees. CONCLUSIONS Our pilot study using NGS for monogenic diabetes screening in the MODY cohort confirmed that it improves the detection of diabetes-related sequence differences. The screening with NGS should also include diabetic patients for whom Sanger-based screening for particular subtypes of MODY provided negative results.
diabetic cases. Thus, we estimate that in Poland the number of patients with MODY is around 20,000 to 30,000. Most cases are usually diagnosed before the age of 25 years; however, many patients are identified in the fourth or fifth decade of life or even later. Some patients with MODY may present with obesity, although it is not a typical feature. Consequently, the overlap of clinical characteristics between MODY, type 1 diabetes, and type 2 diabetes represents a challenge for differential diagnosis. Moreover, the clinical phenotype in individuals with MODY may vary within the same pedigree.

So far around a dozen genes have been reported to be responsible for MODY. Genetic testing has frequently relied on phenotype-guided screening for the most common MODY genes such as HNF1A and GCK. Therefore, molecular screening in many countries, particularly in those where genetic testing is not reimbursed by a national health care system, has been limited only to the 2 or 3 most common MODY genes. Owing to limitations both in funding and access to genetic diagnostic facilities, patients who test negative for mutations in these genes have been underdiagnosed and consequently have often received inappropriate therapy. With the advent of high-throughput next-generation sequencing (NGS) technology, there has been an improvement in screening strategies enabling a simultaneous analysis of a panel of genes at a comparable cost to testing for a few genes by Sanger sequencing.

In this study, we aimed to assess the utility of NGS for detecting mutations in a set of known monogenic diabetes genes, using a cohort of Polish patients with negative results for HNF1A-MODY or GCK-MODY in Sanger sequencing screening.

**PATIENTS AND METHODS**

A contact database of MODY families is maintained at the Department of Metabolic Diseases, Kraków, Poland, based on the following criteria: 1) autosomal dominant inheritance pattern of diabetes mellitus; 2) presence of the disease in at least 3 consecutive generations; 3) at least 2 diabetic family members diagnosed at the age of 30 years or earlier and treated for at least 2 years with diet, oral medication, or insulin at a dose lower than 0.5 U/kg. Details of the ascertainment protocol were described in previous papers.

The current study group consisted of 54 probands who had been screened for mutations in the HNF1A and GCK genes by Sanger sequencing and showed negative results. Additionally, 9 positive controls in whom mutations in the GCK or HNF1A genes were previously identified by Sanger sequencing were included. The study was performed according to the Declaration of Helsinki, and was approved by the Bioethical Committee of the Jagiellonian University.

We performed basic clinical laboratory analyses, including the measurement of hemoglobin A1c (HbA1c), fasting glucose, and C-peptide levels. HbA1c was measured by high-performance liquid chromatography (Bio-Rad, Hercules, California, United States). Fasting glucose levels were determined with the oxidase method. C-peptide levels were determined by an enzyme immunoassay. A diagnosis of diabetic microvascular complications—retinopathy and nephropathy—was established as described previously. Genomic DNA was extracted from peripheral blood with a Maxwell Instrument (Promega, Madison, Wisconsin, United States). Libraries were prepared according to an established protocol.

For target enrichment, we used a custom Agilent SureSelect exon-capture assay with probes designed to target exons and splice sites of 28 genes known to be associated with monogenic forms of diabetes such as MODY, neonatal diabetes, and lipodystrophy. Target genes were as follows: ABCC8, BLK, CEL, E1F2AK3, FOXP3, GATA4, GATA6, GCK, GLIS3, HNF1A, HNF1B, HNF4A, IER3IP1, INS, KCNJ11, KLF11, LMNA, NEUROD1, NEUROG3, PAX4, PDX1, PPARG, PTF1A, RFX6, SLC19A2, SLC2A2, WFS1, and ZFPP7. The final captured library was sequenced on the MiSeq platform (Illumina, San Diego, California, United States) using 75bp paired-end reads.

Data were processed as follows: reads were aligned to the 1000 Genomes Project reference genome (GRCh37-derived reference sequence) using the BWA-MEM 0.7.5a algorithm. Duplicates were removed using SAMTools 0.1.19. Realignment across indels and base quality recalibration were performed with GATK 3.1. Variants were called using a Unified Genotyper and filtered using recommended hard filtering parameters (GATK Best Practices v4). The SnvEff tool was used for variant annotation.

Variants were annotated with chromosome location, genomic coordinate, reference and variant nucleotide, absolute number and percentage of reads containing variant nucleotide, reference and variant amino acid, coding frame, gene name, and overlap with a known single nucleotide polymorphism (SNP) from dbsNP 135. The final variant filtering and prioritization was performed with GEMINI based on alternative allele frequency lower than 0.001 in the European American population in the Exome Sequencing Project and lower than 0.3 in the studied cohort, which was a lower threshold than the one typically used (0.2), owing to the small size of the group and the fact that there was more than 1 examined patient per family. Only variants covered by 8 or more reads and with genotype quality over 19 were included in the final analysis. Variant impacts were divided into severity classes according to GEMINI mapping. Next, the results were sent for confirmation by the Sanger method. Direct sequencing was performed using labeled dieoxy-terminated nucleotides and 1 oligonucleotide primer. A computer analysis of the chromatograms obtained was conducted using the DNA Sequencing Analysis Software (Applied Biosystems, Foster City, California, United States).
First, we confirmed the segregation of 11 mutations detected in the GCK gene. Among them, there were 4 known amino-acid substitutions (Met235Val, Leu164Pro, Leu315His, and Gly175Arg) previously missed in Sanger sequencing and 5 earlier reported sequence differences (Gly318Arg, Gly261Arg, Cys220Ter, Phe438Ser, and Val302Leu) identified in 7 patients screened earlier for the HNF1A mutation. At the time of the examination, 6 of these 11 probands were on diet, 1 was treated with insulin, and 4 were on oral drugs.

The mutation identified in the HNF1A (Ala532Thr) gene was previously described. The proband from this family was earlier diagnosed with gestational diabetes at the age of 37 years along with a relatively mild manifestation of hypoglycemia, which was treated only with diet. She was screened for the GCK mutation at the age of 41 years.

We also positively verified the segregation of the NEUROD1 Arg103Pro mutation. The pedigree, consisting of 17 family members and 11 mutation carriers, has been reported in a separate paper (under review).

<table>
<thead>
<tr>
<th>Proband’s ID number</th>
<th>Sex</th>
<th>Age at diagnosis, y</th>
<th>Previous negative gene screening</th>
<th>Gene with currently identified mutation</th>
<th>Nucleotide change / three letter amino-acid description</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>M007-1 F 29</td>
<td>HNF1A and GCK</td>
<td>GCK</td>
<td>c.703A&gt;G/p.(Met235Val)</td>
<td>Garcia-Herrero et al</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td>M013-1 F 23</td>
<td>HNF1A</td>
<td>HNF4A</td>
<td>c.724G&gt;A/p.(Val242Met)</td>
<td>rs202105574, unpublished, HGMD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M038-1 F 23</td>
<td>GCK</td>
<td>GCK</td>
<td>c.491T&gt;C/p.(Leu164Pro)</td>
<td>rs193922296 likely pathogenic (submitted to ClinVar NCBI, August 18, 2011)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M043-1 M 18</td>
<td>HNF1A</td>
<td>GCK</td>
<td>c.952G&gt;A/p.(Gly318Arg)</td>
<td>Dusatkova et al</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>M048-1 M</td>
<td>HNF1A</td>
<td>GCK</td>
<td>c.781G&gt;A/p.(Gly261Arg)</td>
<td>Beer et al</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>M050-1 F 33</td>
<td>HNF1A</td>
<td>HNF1B</td>
<td>c.1127C&gt;T/p.(Thr376Ile)</td>
<td>not reported previously</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M069-1 F 18</td>
<td>HNF1A and GCK</td>
<td>GCK</td>
<td>c.944T&gt;C/p.(Leu315His)</td>
<td>Dusatkova et al</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>M076-1 F 20</td>
<td>HNF1A</td>
<td>GCK</td>
<td>c.952G&gt;A/p.(Gly318Arg)</td>
<td>Dusatkova et al</td>
<td>10</td>
<td></td>
</tr>
<tr>
<td>M080-1 M 36</td>
<td>HNF1A</td>
<td>GCK</td>
<td>c.660C&gt;G/p.(Cys220Term)</td>
<td>Ziemsse et al</td>
<td>12</td>
<td></td>
</tr>
<tr>
<td>M095-1 F 27</td>
<td>HNF1A</td>
<td>GCK</td>
<td>c.1313T&gt;C/p.(Ph438Ser)</td>
<td>Lorini et al</td>
<td>11</td>
<td></td>
</tr>
<tr>
<td>M103-1 M 36</td>
<td>HNF1A</td>
<td>GLIS3</td>
<td>c.745G&gt;A/p.(Gly249Arg)</td>
<td>not reported previously</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M104-1 M 26</td>
<td>HNF1A</td>
<td>PTF1A</td>
<td>c.75C&gt;T/p.(Asp25Glu)</td>
<td>rs37887768, unpublished, HGMD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M127-1 F 36</td>
<td>GCK</td>
<td>GCK</td>
<td>c.523G&gt;A/p.(Gly175Arg)</td>
<td>Glyn et al</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>M137-1 F 45</td>
<td>HNF1A</td>
<td>HNF4A</td>
<td>c.299G&gt;A/p.(Arg100Gln)</td>
<td>rs37155731, unpublished, HGMD</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M140-1 M 9</td>
<td>HNF1A</td>
<td>GCK</td>
<td>c.904G&gt;C/p.(Val302Leu)</td>
<td>Osbak et al</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M144-1 F 15</td>
<td>HNF1A and GCK</td>
<td>HNF4A</td>
<td>c.401G&gt;A/p.(Arg134Gln)</td>
<td>Harries et al</td>
<td>15</td>
<td></td>
</tr>
<tr>
<td>M158-1 F 45</td>
<td>HNF1A</td>
<td>GCK</td>
<td>c.781G&gt;A/p.(Gly261Arg)</td>
<td>Dusatkova et al</td>
<td>18</td>
<td></td>
</tr>
<tr>
<td>M206-1 F 37</td>
<td>GCK</td>
<td>HNF1A</td>
<td>c.1594G&gt;A/p.(Ala532Thr)</td>
<td>CDSM935976, unpublished, HGMD</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: ID, identification; F, female; HGMD*, The Human Gene Mutation Database; M, male

RESULTS  Sequencing results were generated for all 54 patients and 9 positive controls with a previously identified HNF1A or GCK gene mutation. All previously identified sequence differences (Met57Thr, Leu315His, and Gly318Arg [in 2 patients] in the GCK gene and Gln109* [in 2 patients], Arg238Thr [in 2 patients], and Gly292Argfs*25 in the HNF1A gene) were found by NGS in the control group. High (n = 3) or moderate (n = 76) impact genetic changes in 16 genes were found in the study group.

So far, 21 mutations in 18 patients have been confirmed using the Sanger method (TABLE 1). None of the sequence differences selected for Sanger verification was a false positive finding. Among them, there were 11 mutations in the GCK gene, 1 in HNF1A, 5 in HNF4A, 1 in NEUROD1, 1 in HNF1B, 1 in GLIS3, and 1 in PTF1A (TABLE 1). The segregation of these variants with diabetes has been confirmed in 14 pedigrees.
**DISCUSSION** In the current study, we performed the NGS screening to detect mutations in a set of known monogenic diabetes genes using a cohort of Polish patients with previous negative Sanger sequencing screening results for HNF1A-MODY or GCK-MODY. To our knowledge, this study is the first of this type to investigate the clinical application of NGS in the diagnosis of diabetes in Poland.

Since an accurate etiological diagnosis of monogenic forms of diabetes has been proved to lead to a marked improvement both in patient care and family counseling, the use of cost-efficient, fast, and high-throughput methods for an accurate DNA sequencing is of major medical interest for contemporary diabetology. All mutations previously found by Sanger in our 9 positive controls were also confirmed by the NGS method, which demonstrated the diagnostic accuracy of the latter tool. Furthermore, 4 GCK mutations that had been previously missed by Sanger sequencing were subsequently identified by NGS in our study. It has been reported that NGS is helpful in identifying variants that were earlier missed by traditional sequencing. The MODY database was established almost 20 years ago, and over this period, our samples were screened in several collaborating laboratories in which quality control procedures and certification process have been rigorously introduced. Additionally, our screening with NGS resulted in improved diagnosis of MODY by detecting mutations in genes other
genes that have not been screened owing to the identified mutations were in rare monogenic diabetes.

A similar study conducted in England, which used NGS-based genetic screening of a 29-gene panel implicated in monogenic diabetes, has also shown an increase of 17% in the detection of mutations. In the same study, 6 of 14 newly identified mutations were in rare monogenic diabetes genes that have not been screened owing to the absence of characteristic features of their subtypes. It is noteworthy that the number of patients presenting with monogenic diabetes is currently underestimated. Shields et al reported that more than 80% of MODY cases were undiagnosed. The use of NGS methodology will likely reduce this estimate rapidly. It is expected that this will be associated by a reduction in the financial cost of sequencing. We estimate that the cost of mutation screening in the current set of genes for 1 sample in this study was similar or even lower than the price for 1 gene (consisting of approximately 10 exons) search by the Sanger method. Based on the results assessed by the simulated model, it is expected that generalized screening for monogenic diabetes will become more cost-effective with advances in sequencing technology.

While it is generally believed that findings from NGS should be confirmed by Sanger sequencing, this confirmation involves only specific exons with suspected sequence differences. Moreover, it is expected that NGS technology will become increasingly reliable with time. For example, the US Food and Drug Administration has already approved the MiSeq platform for cystic fibrosis mutation screening (http://www.illumina.com/clinical/diagnostics/miseqdx-instrument.html).

There are also reports on other monogenic diseases, which show that targeted NGS sequencing is currently as reliable as the Sanger method. In line with this, none of the variants selected by us for Sanger verification were false positive, unlike some earlier data generated by NGS, which contained such errors. However, those previous data came from whole-genome sequencing research and included low-coverage data. Our research involved targeted sequencing in a relatively low number of genes with higher coverage. Additionally, we used a newer platform with high reliability. We also focused on the search for rare mutations in a limited cohort of probands from highly selected families.

In conclusion, we demonstrated the feasibility of using NGS for monogenic diabetes screening in routine genetic testing of our MODY cohort in...
this pilot study. The screening should include patients with suspected MODY that had not been previously confirmed by Sanger-based screening.

Contribution statement MS was responsible for study design, protocol development, acquisition and interpretation of data, drafting the article, and writing the manuscript; ALG—for study design, acquisition, interpretation of data; FR—for acquisition and interpretation of data; JS—for study design and protocol development; BZ and TP—for acquisition and interpretation of data; TK—for acquisition of data; BKW, MB, and WM—for interpretation of data; FW—for revising the manuscript for important intellectual content and interpretation of data; MTM—for study design, project coordination, writing the manuscript, and approval of the final version of the manuscript.

Acknowledgments We would like to thank the subjects who participated in the study. We are also grateful to Ms. Julita Machlowska and Ms. Agata Porębska for technical assistance. We also acknowledge the editorial assistance of Mr. Cyrus M. Sani, School of Medicine in English, Jagiellonian University Medical College.

The study was funded by a grant from the Polish Ministry of Science and Higher Education: “Iuventus Plus”, No. 0545 / IP1 / 2011/71, “Search and clinical characteristics of the rare forms of MODY” (granted to MS). The Omicron laboratory where the NGS experiments were performed was funded from European Union funds (7th Framework Programme; Call number, FP7-REG-POT-2011-1; Grant number, 286038).

REFERENCES

Badania genetyczne w kierunku cukrzycy monogenowej za pomocą metody sekwencjonowania nowej generacji u pacjentów z cukrzycą typu MODY

Magdalena Szopa1,2*, Agnieszka Ludwig-Gałęzowska1,3*, Piotr Radkowski3, Jan Skupień1,3, Barbara Zapała4, Teresa Platek4, Tomasz Klupa1,2, Beata Kiec-Wilk1,2, Maciej Borowiec5,6, Wojciech Młynarski6, Paweł Wołkowski7, Maciej T. Małecki1,2

1 Katedra Chorób Metabolicznych, Uniwersytet Jagielloński, Collegium Medicum, Kraków
2 Oddział Kliniczny Kliniki Chorób Metabolicznych, Szpital Uniwersytecki w Krakowie, Kraków
3 OMICRON, Uniwersytet Jagielloński, Collegium Medicum, Kraków
4 Katedra Biochemii Klinicznej, Uniwersytet Jagielloński, Collegium Medicum, Kraków
5 Katedra Genetyki Klinicznej i Laboratatoryjnej, Uniwersytet Medyczny w Łodzi, Łódź
6 Klinika Pediatrii, Onkologii, Hematologii i Diabetologii, I Katedra Pediatrii Uniwersytet Medyczny w Łodzi, Łódź

SŁOWA KLUCZOWE
cukrzyca
monogenowa,
sekwencjonowanie

STRESZCZENIE

WPRAWODZENIE Diagnostyka molekularna cukrzycy monogenowej jest istotna w ramach zindywidualizowanej opieki nad pacjentem. Sekwencjonowanie nowej generacji (next-generation sequencing – NGS) umożliwia jednoczesną analizę wielu genów w jednym teście.

CELE Celem badania była ocena możliwości wykorzystania metody NGS do wykrywania znanych mutacji genów związanych z cukrzycą monogenową w grupie polskich pacjentów z cukrzycą typu MODY (maturity-onset diabetes of the young) z uzyskanym wcześniej za pomocą sekwencjonowania sangerowskiego negatywnym wynikiem w kierunku HNF1A-MODY lub GCK-MODY.

PACJENCI I METODY Wybrano panel 28 genów chromosomalnych, w których opisano mutacje będące przyczyną cukrzyc monogenowych. NGS wykonano przy użyciu platformy MiSeq. Zaprojektowano zestaw sond, które pokrywały całe regiony eksomowe i regiony splice site. Do badania wybrano 54 pacjentów z wcześniejszym negatywnym wynikiem w kierunku mutacji w genie HNF1A lub GCK uzyskanym metodą Sangera.

WYNIKI Otrzymano wyniki NGS u wszystkich 54 pacjentów i 9 pozytywnych kontroli z wcześniejszej zidentyfikowanych mutacji genów związanych z cukrzycą monogenową w grupie polskich pacjentów. Wśród 28 analizowanych genów mutacje wykryto w 16. Charakter zmian genetycznych włączonych do analizy został opisany w badaniu NGS jako wysoki (n = 3) lub umiarkowany (n = 76). Wśród wykrytych mutacji zidentyfikowano też 4 znane mutacje w genie GCK, które nie zostały wcześniej wykryte w badaniu metodą Sangera. Jak dotąd potwierdzono sekwencjonowaniem sangerowskim 21 mutacji wykrytych za pomocą NGS, a segregację z cukrzycą w 14 rodzinach.

WNIOSKI Nasze badanie pilotażowe z wykorzystaniem NGS, jako metody przesiewowej w diagnostyce cukrzycy monogenowej w kohorcie MODY potwierdziło jej przydatność w lepszym wykrywaniu różnic w sekwencjach genetycznych związanych z cukrzycą. Badania przesiewowe z wykorzystaniem NGS powinny także obejmować pacjentów z cukrzycą, u których wcześniejsze sekwencjonowanie sangerowskie w kierunku pojedynczych podtypów MODY daly negatywne wyniki.