Co-segregation analysis and functional trial in vivo of candidate genes for monogenic diabetes

Ingrida Stankute, Mintaute Kazlauskiene, Jean-Louis Blouin, Valerie M Schwitzgebel, Rasa Verkauskiene

ABSTRACT

Introduction The aim of this study was to perform familial co-segregation analysis and functional trial in vivo during mixed meal tolerance test (MMTT) of novel variants in diabetes candidate genes.

Research design and methods It is a continuation of the project “Genetic diabetes in Lithuania” with the cohort of 1209 patients with diabetes. Prior screening for autoimmune markers confirmed type 1 diabetes (T1D) diagnosis in 88.1% (n=1065) of patients, and targeted next-generation sequencing identified 3.5% (n=42) pathogenic variants in MODY genes. Subsequently, 102 patients were classified as having diabetes of unknown etiology. 12/102 were found to have novel variants in potential diabetes genes (RX2, RREB1, SLC5A1) with variants in RFX2, MCHR, CASP10, TMPRSS6, HGFC, DACH1, ZBED3. Co-segregation analysis and MMTT were carried out in order to study beta-cell function in subjects with specific variants.

Results MMTT analysis showed that probands with variants in MC4R, CASP10, TMPRSS6, HGFC, and SLC5A1 (c.1415T>C) had sufficient residual beta-cell function with stimulated C-peptide (CP) >500 pmol/L. Seven individuals with variants in RX2, RREB1, GCKR, DACH1, ZBED3 and SLC5A1 (c.1415T>C, c.392A>T) presented with complete beta-cell failure. No statistical differences were found between patients with sufficient CP production and those with complete beta-cell failure when comparing age at the onset and duration of diabetes. Nineteen family members were included in co-segregation analysis; no diabetes cases were reported among them. Only in patient with the variant c.1894G>A in RFX2 gene, none of the family members were affected by proband’s variant.

Conclusions Functional beta-cell study in vivo allowed to select five most probable genes for monogenic diabetes. Familial co-segregation analysis showed that novel variant in RFX2 gene could be a possible cause of diabetes. Future functional analysis in vitro is necessary to support or rule out the genetic background as a cause of diabetes.

WHAT IS ALREADY KNOWN ON THIS TOPIC
⇒ To date, more than 40 genes are identified as causative for monogenic diabetes. However, despite of the increasing usage of genetic testing, up to 10% of cases remain with unknown etiology of diabetes.

WHAT THIS STUDY ADDS
⇒ Our study revealed novel variants in 10 diabetes candidate genes (RFX2, RREB1, SLC5A1, GCKR, MCHR, CASP10, TMPRSS6, HGFC, DACH1, ZBED3).

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY
⇒ Future in vitro experiments are necessary to support or exclude these candidate genes as cause of diabetes.

INTRODUCTION

Since the discovery of the common causes of monogenic diabetes (MD), genetic diabetes research brought remarkable progress during the last three decades. To date, more than 40 different genetic subtypes of MD have been identified, each having specific phenotype and pattern of inheritance. However, the wide usage of next-generation sequencing (NGS) delivered many novel variants in already known and potential diabetes genes.

Positive autoimmune markers are the hallmark of type 1 diabetes (T1D) diagnosis. Glutamic acid decarboxylase autoantibodies (GADA), insulinoma-associated antigen 2 autoantibodies (IA-2A), and insulin autoantibodies (IAA) are the most commonly used biomarkers to confirm T1D and rule out patients from unnecessary genetic testing. The implementation of zinc transporter 8 (ZnT8) autoantibodies to clinical practice improved discrimination of T1D from other forms of diabetes due to the high expression of ZnT8 in pancreatic beta cells.

Despite advances in genetic testing and discovery of new biomarkers, up to 10% of patients remain in gray area with still unknown etiology of diabetes.

The present study is a continuation of Lithuanian-Swiss project “Genetic Diabetes in Lithuania.” Data on study cohort are published previously; briefly, the project’s the cohort included 1209 patients with diabetes. Prior screening for autoimmune
markers confirmed T1D diagnosis in 88.1% (n=1065) of patients, and targeted next-generation sequencing identified 3.5% (n=42) pathogenic variants in known MODY genes, and 1% (n=12) of patients were found to have novel variants in potential diabetes genes.

The aim of this study was to perform familial co-segregation analysis and functional pancreatic beta-cell studies in vivo during mixed meal tolerance test (MMTT) for patients with novel variants in potential diabetes genes in order to assess whether the specific variants may determine the clinical phenotype.

**RESEARCH DESIGN AND METHODS**

**Study patients and design**

The study recruited 102 young patients with unknown etiology diabetes (63.7% (n=65) children <18 years, 36.3% adults 18–25 years); 53.9% (n=55) were males. The whole design of the project is presented in online supplemental figure 1.

**Novel variants in potential diabetes genes**

Genetic analysis was carried out using high-throughput sequencing of 307 genes related to diabetes, previously described in detail. In total, 12/102 (11.8%) were identified as having new likely pathogenic (class 4) variants, according to American College of Medical Genetics and Genomics guidelines. The variants were found in the following genes: DACH1, RREB1, ZBED3, and RFX2, which encode transcription factors; SLC5A1 encoding sodium-glucose cotransporter; TMPRSS6 encoding a protein that is mainly involved in iron balance in the body; CASP10 encoding a protein of caspase family, which plays a key role in execution of cell apoptosis; the product of GCKR gene inhibiting glucokinase in liver and pancreatic cells; HGFAC encoding a hepatocyte growth factor activator; MC4R encoding membrane-bound receptor and member of the melanocortin receptor family responsible for metabolism and appetite.

**Recruitment of family members, their clinical assessment, and co-segregation analysis**

Overall, 19 family members of 7 patients with novel variants were recruited into the study (2 patients were siblings); the remaining families did not agree to participate in the study. All first-line relatives of seven patients were included, both parents and all siblings of each patient. Data for pedigree analysis, clinical information, and blood samples for lab analysis and DNA extraction were collected during the first appointment.

Clinical data about general health, medical history, physical examination including height, weight, body mass index (BMI), and blood pressure were evaluated for all family members. To assess beta-cell function and metabolism of carbohydrates, oral glucose tolerance test (OGTT) was performed and evaluated following ISPAD guidelines for children and ADA protocol for adults; for the patient under 5 years of age, blood glucose (BG), fasting insulin, C-peptide (CP), and HbA1c were evaluated in the fasting state. HbA1c and autoimmunity status (GADA, IA-2A, IAA, ZnT8A) were evaluated for all members, following previously described laboratory assays. ZnT8A were measured by ELISA (BioVendor, Czech Republic) following manufacturer’s guidelines, and the value ≥15 U/mL was assigned as positive. The assay kit was able to detect and quantify all three isoforms (ZnT8R325, ZnT8W325, ZnT8Q325) of ZnT8 antibodies.

**Variant confirmation**

Putative pathogenic point variants identified by next-generation sequencing (NGS) were confirmed by PCR and Sanger sequencing using forward (F) and reverse (R) primers for variant confirmation.

<table>
<thead>
<tr>
<th>Gene</th>
<th>Exon</th>
<th>Primer ID</th>
<th>Sequence</th>
<th>MIX</th>
</tr>
</thead>
<tbody>
<tr>
<td>DACH1</td>
<td>1</td>
<td>DACH1-x1F</td>
<td>CCGCTTCAGCTTGGTGTTGTA</td>
<td>KAPA2G-Robust</td>
</tr>
<tr>
<td>HGFAC</td>
<td>6</td>
<td>HGFAC-x6F</td>
<td>CTGCTCAAATCCAGGACC</td>
<td>KAPA2G</td>
</tr>
<tr>
<td>RREB1</td>
<td>10</td>
<td>RREB1-x10F</td>
<td>TCCATCCCCAAAGAAGTCCAG</td>
<td>KAPA2G-Robust</td>
</tr>
<tr>
<td>RFX2</td>
<td>15</td>
<td>RFX2-x15F</td>
<td>TGCAGTCTTACAGCAGTCA</td>
<td>KAPA2G</td>
</tr>
<tr>
<td>SLC5A1</td>
<td>9</td>
<td>SLC5A1-x9F</td>
<td>GGAACCTCGTCTGCTGTTGG</td>
<td>KAPA2G-Robust</td>
</tr>
<tr>
<td>SLC5A1</td>
<td>12</td>
<td>SLC5A1-x12F</td>
<td>AGGTGTTTTATTCTCTGGTCT</td>
<td>KAPA2G</td>
</tr>
<tr>
<td>GCKR</td>
<td>17</td>
<td>GCKR-x17F</td>
<td>ACATGTGTTTTCTGTGTCA</td>
<td>KAPA2G-Robust</td>
</tr>
<tr>
<td>GCKR</td>
<td>17</td>
<td>GCKR-x17R</td>
<td>AAGCACAAGAAAAGGAGC</td>
<td>KAPA2G-Robust</td>
</tr>
</tbody>
</table>
primes described in table 1. PCR was using mix KAPA2G or KAPA2G_robust reagent at manufacturer conditions (Sigma-Aldrich), and amplification cycles performed on Biometra TOne thermal cycler.

Sequencing was performed using BigDye terminator at manufacturer conditions (Thermo Fisher) and products analyzed with Abi 3500 genetic analyzer (Thermo Fisher). Sequencing data were then analyzed using Mutation Surveyor (Softgenetics), using GenBank references.

Evaluation of beta-cell function of carriers of specific variants in vivo with MMTT

MMTT with nine time points (−10′, 0′, 15′, 30′, 60′, 90′, 120′, 150′, 180′) was performed for all patients with novel variants, following standardized protocol, described previously.21 CP and BG levels were measured at all time points. Area under the curve of CP (AUC_{CP}), the baseline (CP_{Base}), peak (CP_{max}), and average (CP_{ave}) levels of CP were analyzed, as well as the kinetics of BG levels. In the absence of published data about our novel variants and their potential effects on beta-cell function, stimulated CP level <200 pmol/L has been used as a threshold for complete beta-cell failure, as indicated by several authors previously.22–25

Statistical analysis

SPSS V.23.0 software package was used for statistical analysis. The median (min; max) values were used for representation of the data, unless indicated otherwise. The Mann-Whitney U test was used for comparisons of non-normally distributed data, and the χ² test was used for categorical variables. P value <0.05 was considered as statistically significant, all tests were two-tailed.

RESULTS

Characteristics of the cohort

General characteristics of the cohort are presented in table 2.

Clinical data and beta-cell function analysis of probands with novel genetic variants

Twelve subjects were identified to have novel variants in potential diabetes genes, overall, 11 different variants in 10 candidate genes. Moreover, 83.3% (n=10) were males; the median age of the group was 199 months (82; 315).

### Table 2 General characteristics of the cohort (n=102)

| Characteristics                  | Value
|----------------------------------|-------
| Age at diabetes Dx, months       | 92 (8; 292) |
| BG at diabetes Dx, mmol/L        | 16.8 (4; 39) |
| CP at diabetes Dx, pmol/L        | 350 (30; 4510) |
| Ketosis at diabetes Dx, % (n)    | 64.2 (61) |
| Diabetes in family history, % (n)| 43.1 (44) |
| Age at analysis, months          | 192 (19; 315) |
| BMI Z-score at analysis          | 0.0 (−2.9; 2.6) |
| HbA1c at analysis, % mmol/mol    | 8.4 (4.8; 14.5) |
| Treatment at investigation, % (n)| 89.2 (91) |
| Novel variants in potential diabetes genes, % (n) | 11.8 (12) |

Data presented as median (min; max), unless stated otherwise. Dx, diagnosis; BG, blood glucose; CP, C-peptide; BMI, body mass index; HbA1c, glycosylated haemoglobin.

### Table 3 Kinetics of C-peptide and blood glucose during MMTT in probands with novel variants

<table>
<thead>
<tr>
<th>Affected gene</th>
<th>CP kinetics</th>
<th>BG kinetics</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AUC, pmol/L/180 min</td>
<td>CP_{Base}, pmol/L</td>
</tr>
<tr>
<td></td>
<td>AUC, mmol/L/180 min</td>
<td>BG_{Base}, mmol/L</td>
</tr>
<tr>
<td>/RFX2</td>
<td>81</td>
<td>1</td>
</tr>
<tr>
<td>/RREB1</td>
<td>580</td>
<td>21</td>
</tr>
<tr>
<td>/SLC5A1</td>
<td>44</td>
<td>3</td>
</tr>
<tr>
<td>/GCKR</td>
<td>383</td>
<td>38</td>
</tr>
<tr>
<td>/MC4R</td>
<td>3702</td>
<td>287</td>
</tr>
<tr>
<td>/CASP10</td>
<td>2952</td>
<td>311</td>
</tr>
<tr>
<td>/TMPPRS6</td>
<td>1159</td>
<td>52</td>
</tr>
<tr>
<td>/HGFAC</td>
<td>4432</td>
<td>302</td>
</tr>
<tr>
<td>/DACH1</td>
<td>37</td>
<td>2</td>
</tr>
<tr>
<td>/SLC5A1*</td>
<td>4201</td>
<td>253</td>
</tr>
<tr>
<td>/ZBED3</td>
<td>42</td>
<td>1</td>
</tr>
<tr>
<td>/SLC5A1*</td>
<td>256</td>
<td>17</td>
</tr>
</tbody>
</table>

Probands with CP_{max} >200 pmol/L are in bold.

*The same DNA change in SLC5A1 gene, probands are siblings.

AUC, area under the curve; BG, blood glucose; BG_{max}, the average of blood glucose; BG_{Base}, the baseline level of blood glucose; BG_{ave}, the peak of blood glucose; CP, C-peptide; CP_{ave}, the average of C-peptide; CP_{Base}, the baseline level of C-peptide; CP_{max}, the peak of C-peptide; MMTT, mixed meal tolerance test.
The median age at diabetes onset was 117 months (24; 214), and 83.3% (n=10) presented with ketosis, and the median BG at the onset was 15.8 mmol/L (6.2; 32). At investigation, the median HbA1c was 8.1% ((5.8; 11.6) 65 mmol/mol (39.9; 103.3)), BMI Z-score −0.98 (−1.7; 2.4).

Eleven patients (91.7%) were treated with insulin, with the median dose 0.6 U/kg/day (0.2; 1.1); one patient had no pharmacological treatment. All probands were negative for GADA, IA-2A; 58.3% (n=7) were positive for IAA, identified after introduction of insulin injections; one patient was positive for ZnT8 autoantibodies, but had strong clinical suspicion of MD. Clinical and genetic data for each patient are provided in online supplemental table 1.

MMTT data analysis showed that five probands exceeded the threshold of stimulated CP (200 pmol/L) during MMTT: no. 5 with affected gene MC4R; no. 6, CASP10; no. 7, TMPRSS6; no. 8, HGFAC; and no. 10, SLC5A1. Data on kinetics of CP and BG during MMTT for each proband is presented in table 3. All CP and BG responses during MMTT are presented in figure 1. There were no statistical differences when comparing age at the onset and diabetes duration between patients with stimulated CP >200 pmol/L versus patients with stimulated CP <200 pmol/L.

**Co-Segregation and pedigree analysis**

Nineteen members of six families (seven patients) were analyzed. Twelve of them were parents, seven siblings; there were no offspring of probands. Neither family member presented with clinical signs of diabetes or was found to have diabetes by OGTT or HbA1c. None of the family members were identified as positive for diabetes autoimmune markers. Father of proband no. 1 had impaired fasting glycemia 6.17 mmol/L. The sibling of proband no. 3 was positive for IAA with HbA1c 6.5% (48 mmol/mol) and normal OGTT. The sibling of proband no. 9 with normal BMI had impaired glucose tolerance (IGT)—7.83 mmol/L. Co-segregation analysis revealed single proband (no. 1) with the novel variant in RFX2 gene whose family members did not carry the variant. The pedigree of each family and glucose metabolism parameters are shown in figure 2.

**DISCUSSION**

The main strength of the whole project “Genetic Diabetes in Lithuania” is that it included the entire cohort of children (n=860) with diabetes and in the majority of young diabetic adults (n=349) in Lithuania, and it reflects our approach to the search for genetic diabetes cases. During this project, 3.5% (n=42) of subjects were identified to have MD caused by pathogenic variants in already known MD genes.7 Even after testing for pancreatic beta-cell autoimmune markers, and performing comprehensive targeted NGS, in 8.4% (102/1209) of the total cohort precise diabetes etiology was not identified. As expected, high-throughput sequencing brought several novel variants (11/102) in potential diabetes genes, in agreement with other authors reporting candidate gene variants in up to 10% of sequenced cohorts.26-28 Genomic individualized medicine not only brings the accurate molecular diagnosis and the most effective treatment plan for the patient but also delivers surveillance plan for family members; furthermore, it has proven cost-effectiveness for the healthcare system. Therefore, with this study, we also present our selection of potential genes for MD for further functional analysis in vitro.
although it requires time and effort. Thus, the functional trial of beta cells in vivo during MMTT in our study showed that even after more than 2.5 years of diabetes duration, patients with novel variants in MC4R, CASP10, TMPRSS6, HGFAC, and SLC5A1 (one of brothers with c.1415T>C) still had sufficient residual pancreatic function, suggesting either slower beta-cell destruction or stable functional residual beta-cell mass. Furthermore, the rest of the subjects, who had variants in RFX2, RREB1, GCKR, DACH1, and ZBED3, also in SLC5A1, showed complete beta-cell failure and endogenous insulin deficiency. We believe that hyperglycemia correction with fast-acting insulin 2 hours prior to MMTT in subjects with DACH1 gene variant could influence glycemia levels until the 90th minute, considering C-peptide levels.

Considering that several previous studies have shown that serum CP, fasting or random, and urinary CP/creatinine ratio are highly sensitive and specific biomarkers for discriminating T1D and other forms of diabetes, especially monogenic, we believe that our results suggest that patients with sufficient endogenous insulin production may be eligible for treatment optimization in the future, given that our previous treatment optimization trial in patients with maturity-onset diabetes of the young (MODY) showed the threshold of stimulated CP ≥332.5 pmol/L for a successful treatment transition (from insulin injections to oral therapy).

Additional strength of our study is familial co-segregation analysis. This analysis linked the novel variants in candidate genes with diabetes phenotype. It revealed that the variant c.1894G>A in RFX2 gene could be a suitable candidate for functional analysis in vitro, as neither of healthy family members possessed the novel variant of the proband. Despite the positivity for ZnT8A, this proband was intentionally selected for NGS because of clinical suspicion of MD and data reported by other authors that up to 1–2% of patients with MODY are found to have positive autoimmune markers. Further genetic analysis revealed that the patient had a novel variant. RFX2 gene belongs to the regulatory factor binding to the X-box.
(RFX) transcription factors’ family.\(^3\) This family also includes RFX6 gene, which is a known beta-cell transcription factor and causes permanent neonatal diabetes, or MODY.\(^4\)\(^5\) Usually, pathogenic mutation in RFX6 presents with hypoplasia of pancreas and additional syndromic features, although none of them were reported in our patient. Moreover, the recent study in animal models showed that alterations in the expression of RFX2 gene may lead to increased susceptibility for autoimmune insulitis in beta cells.\(^3\)\(^4\) Taking into account all data, we believe that RFX2 could be a good candidate for functional analysis in cell cultures.

The main limitation of this study was the drop-out of patients due to cumbersome and time-consuming investigations. Furthermore, COVID-19 pandemic prolonged the recruitment of study subjects for a couple of years. It is reported by several authors that the willingness to participate in clinical trials during COVID-19 pandemic decreased up to 10–20%; some were closed because of insufficient sample sizes.\(^3\)\(^5\)\(^6\) Nevertheless, further investigations of the remaining families of our cohort are to be implemented. An additional limitation of our study is that we have not yet been able to reclassify novel genetic variants; therefore, we need future functional experiments of candidate genes in vitro.

In summary, functional beta-cell study in vivo allowed to select five most probable diabetes genes (MCAR, CASP10, TMPRSS6, HGFAC, SLC5A1) after proven endogenous insulin secretion during MMTT. Changes in RFX2 gene may have either a causative or predisposing role in diabetes pathophysiology. Further research directions with functional analysis in cell cultures of these genes could support or rule out the genetic background as a reason of diabetes.

**Contributors** IS participated in the conception and idea of the study, subjects’ enrollment, data collection, database creation, carried out statistical analyses and wrote the manuscript. MK performed laboratory analyses. J-LB performed genetic analyses of participants. VMS participated in the design and coordination of the study, and reviewed and edited the manuscript. VMS and RV equally contributed to the study. IS takes full responsibility as a guarantor for the integrity, and accuracy of the data, and publishment process.

**Funding** The study was supported by a grant from Lithuanian Research Council Lithuanian-Swiss program “Research and development” (CH-3-SMM-01/09) and the Federal Department of Foreign Affairs of Switzerland. Swiss National Science Foundation Grant No. CR3333 140655 and CR3333 1166591 to VMS.

**Competing interests** None declared.

**Patient consent for publication** Not applicable.

**Ethics approval** This study involves human participants, and regional and national ethical approvals from Kaunas Regional Biomedical Research Ethics Committee and Lithuanian Bioethics Committee (No. BE-2-5 and No. BE-2-51) were obtained for this study. Participants were included in the study after they and/or their official guardians gave an informed consent. The study was conducted following the Declaration of Helsinki.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available on reasonable request.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iD** Ingrida Stankute http://orcid.org/0000-0002-0228-7143

**REFERENCES**


