diabetesresearchandclinicalpractice185(2022)109226

Contents lists available at ScienceDirect

Diabetes Research and Clinical Practice

journal homepage: www.journals.elsevier.com/diabetes-research-and-clinical-practice

Fucosylated AGP glycopeptides as biomarkers of HNF1A-Maturity onset diabetes of the young

Marko Tijardović [a], Tamara Štambuk [b], Agata Juszczak [c,d], Toma Keser [a], Daniela Gasperikova [e], Mislav Novokmet [b], Erling Tjora [f], Edita Pape Medvidović [g], Juraj Stanik [e,h], Pål Rasmus Njølstad [f,i], Gordan Lauc [a,b], Katharine R. Owen [c,d], Olga Gornik [a,*]

[a] Faculty of Pharmacy and Biochemistry, University of Zagreb, Zagreb, Croatia
[b] Genos Glycoscience Research Laboratory, Zagreb, Croatia
[c] Oxford Centre for Diabetes, Endocrinology and Metabolism, Churchill Hospital, University of Oxford, Oxford, UK
[d] Oxford NIHR Biomedical Research Centre, Oxford University Hospitals NHS Foundation Trust, John Radcliffe Hospital, Oxford, UK
[e] Department of Metabolic Disorders, Institute of Experimental Endocrinology, Biomedical Research Center, Slovak Academy of Sciences, Bratislava, Slovakia
[f] Children and Youth Clinic, Haukeland University Hospital, Bergen, Norway
[g] Vuk Vrhovac University Clinic for Diabetes, Endocrinology and Metabolic Diseases, Merkur University Hospital, Zagreb University School of Medicine, Zagreb, Croatia
[h] Department of Pediatrics, Medical Faculty of Comenius University and National Institute for Children's Diseases, Bratislava, Slovakia
[i] Center for Diabetes Research, Department of Clinical Science, University of Bergen, Bergen, Norway

ARTICLE INFO

Keywords:
MODY
Type 2 diabetes mellitus
Alpha-1-acid glycoprotein
N-glycosylation
Glycoprotein
LC-MS

ABSTRACT

Aims: We previously demonstrated that antennary fucosylated N-glycans on plasma proteins are regulated by HNF1A and can identify cases of Maturity-Onset Diabetes of the Young caused by HNF1A variants (HNF1A-MODY). Based on literature data, we further postulated that N-glycans with best diagnostic value mostly originate from alpha-1-acid glycoprotein (AGP). In this study we analyzed fucosylation of AGP in subjects with HNF1A-MODY and other types of diabetes aiming to evaluate its diagnostic potential.
Methods: A recently developed LC-MS method for AGP N-glycopeptide analysis was utilized in two independent cohorts: a) 466 subjects with different diabetes subtypes to test the fucosylation differences, b) 98 selected individuals to test the discriminative potential for pathogenic HNF1A variants.
Results: Our results showed significant reduction in AGP fucosylation associated to HNF1A-MODY when compared to other diabetes subtypes. Additionally, ROC curve analysis confirmed significant discriminatory potential of individual fucosylated AGP glycopeptides, where the best performing glycopeptide had an AUC of 0.94 (95% CI 0.90–0.99).
Conclusions: A glycopeptide based diagnostic tool would be beneficial for patient stratification by providing information about the functionality of HNF1A. It could assist the interpretation of DNA sequencing results and be a useful addition to the differential diagnostic process.

1. Introduction

Maturity-onset diabetes of the young caused by *HNF1A* variants (HNF1A-MODY) is the most common type of monogenic diabetes in adults. Monogenic diabetes, unlike more common diabetes types, is characterized by single gene disorder, usually resulting in functional defects of pancreatic beta-cells that cause moderate to severe hyperglycaemia [1]. HNF1A-MODY is caused by pathogenic variants in *HNF1A*, a gene coding for a transcription factor present in many tissues, resulting in progressive beta-cell dysfunction and hyperglycemia. Typical features include young age of onset, autosomal dominant inheritance, absence of autoimmunity and reduction in beta-cell insulin secretion [2].

Being a rare type of diabetes and having clinical similarities with other more common types, HNF1A-MODY is often misdiagnosed. Its clinical features overlap with both type 1 (T1DM) and type 2 (T2DM) diabetes, resulting in around 80% of MODY cases initially being incorrectly diagnosed as T1DM or T2DM and therefore leading to significant delays in proper diagnosis and treatment [3,4].

The first-line treatment for HNF1A-MODY [5] is low dose of

* Corresponding author at: Faculty of Pharmacy and Biochemistry, University of Zagreb, Ante Kovačića 1, 10000 Zagreb, Croatia.
E-mail address: ogornik@pharma.unizg.hr (O. Gornik).

https://doi.org/10.1016/j.diabres.2022.109226
Received 9 December 2021; Received in revised form 19 January 2022; Accepted 28 January 2022
Available online 2 February 2022

sulfonylureas providing improved glycemic control and better quality of life compared to other treatments like insulin [6]. Furthermore, autosomal dominant inheritance, along with high genetic penetrance result in high probability for family members to also be affected or at risk of developing MODY [7,8], further confirming the importance of correct and timely diagnosis.

Confirmative diagnosis of MODY is direct sequencing of the genes related to suspected MODY type (e.g., *HNF1A*) but in some cases the genetic variant is difficult to interpret. If the encountered genetic variant is not previously characterized, further assessment of pathogenicity is needed, sometimes leading to inconclusive results and therefore inability to confirm or exclude MODY diagnosis [9].

N-glycosylation is a co- and post-translational modification of proteins, characterized by addition of oligosaccharide attachments, present on almost every plasma protein [10]. Changes in plasma protein N-glycosylation have been associated with various diabetes subtypes in many studies [11–15]. Our GWAS of the plasma protein N-glycome identified *HNF1A* as a key regulator of plasma protein fucosylation, a process of fucose addition to glycan structure, as a part of the N-glycosylation process [16]. Based on this finding, we further showed that N-glycans in individuals with HNF1A-MODY have reduced levels of plasma protein fucosylation compared to other types of diabetes or healthy controls [17]. Moreover, our previous studies confirmed plasma protein N-glycosylation to be not only a powerful differentiator of HNF1A-MODY from common forms of diabetes (the best performing N-glycan had an AUC of 0.90), but also associated with the predicted effect on function of the *HNF1A* variant being assessed. Thus we hypothesized that N-glycosylation was a further tool to identify pathogenic variants [18] and could therefore assist diagnostic decisions in cases of unclear *HNF1A* variant assessment results. Besides, with genetic testing not being readily available in some countries, there is additional value in development of non-genetic MODY diagnostic tools [3].

Plasma N-glycosylation analysis presents a cumulative analysis of all N-glycans present in a plasma sample of an individual. Even though directly measuring the functionality of *HNF1A* gene through total plasma N-glycome profiling represents a promising approach, it also has some drawbacks. Not all plasma proteins are expected to be equally affected by the downregulation of fucosylation as a result of *HNF1A* mutation even though all of them contribute to the measured plasma profile. More importantly, the abundance of individual glycan structures within the plasma glycoprofile is also dependent on the relative plasma concentration of glycoproteins carrying these glycans, consequently introducing variation and uncertainty into obtained measurements.

Both studies that analyzed plasma N-glycosylation in HNF1A-MODY patients have found complex, triantennary N-glycans to be the most affected glycans in terms of fucosylation reduction in HNF1A-MODY patients and therefore the most informative structures for diagnostic differentiation. These glycan structures are known to predominantly originate from alpha-1-acid glycoprotein (AGP) [19], a heavily glycosylated acute phase plasma protein [20] suggesting its altered N-glycosylation in MODY patients.

An AGP molecule has 5 glycosylation sites which can all carry different N-glycans [21]. Since two isoforms of AGP, differing in some glycosylation sites, are present in human plasma, a total of 8 glycopeptides can be obtained by its tryptic digestion [22]. Recently we have developed and described a high-throughput method for site-specific N-glycosylation analysis of AGP [23].

We applied this method to characterize the changes in N-glycosylation of AGP in HNF1A-MODY compared to other diabetes types (T1DM, T2DM, HNF4A-MODY). The aim was to identify specific glycosylation changes and individual glycopeptides differing in HNF1A-MODY subjects. Their ability to effectively discriminate subjects with HNF1A-MODY from other young adult-onset non-autoimmune diabetes subjects was also assessed.

2. Subjects

2.1. Participants with HNF1A-MODY, HNF4A-MODY, T1DM and T2DM

In order to identify differently glycosylated AGP peptides and to test if the changes are unique to HNF1A-MODY, the analysis was performed on participants with various diabetes types. The cohort consisted of 466 participants grouped by the type of diabetes. It included 109 with HNF1A-MODY, 17 with HNF4A-MODY, 69 with T1DM, and 271 with T2DM. The patients were recruited from Norway (Norwegian MODY registry), Slovakia (diabetes outpatient clinics throughout Slovakia), and Oxford, UK (Young Diabetes in Oxford (YDX) study). Participants with HNF1A- and HNF4A-MODY had diagnosis confirmed by sequencing in a certified diagnostic center, while participants with type 1 and type 2 diabetes were not sequenced. Type 1 diabetes was diagnosed based on C-peptide levels < 0.1 nmol/L, and/or positive glutamic acid decarboxylase (GAD) antibodies, with permanent insulin therapy from time of diagnosis. Individuals with type 2 diabetes were diagnosed before 45 years of age, had absent GAD antibodies, C-peptide > 0.2 nmol/l, and were without requirement for permanent insulin within 3 months of diagnosis. Misdiagnosed MODY cases potentially present in this group are not expected to confound the comparison due to low incidence of MODY in relation to type 2 diabetes. The study was approved by corresponding ethics committees and all participants gave written informed consent.

2.2. Participants with young adult-onset non-autoimmune diabetes

This cohort consisted of participants with young adult-onset non-autoimmune diabetes, a group with elevated risk for HNF1A-MODY. It was studied to additionally confirm observations from the first part of the study and to test the ability of AGP N-glycopeptides to discriminate subjects with HNF1A-MODY from other young adult-onset non-autoimmune diabetes subjects. The cohort included 98 individuals from UK and Croatia. The UK participants ($N = 50$) were recruited through the Young Diabetes in Oxford (YDX) study while Croatian participants ($N = 48$) were recruited through the Croatian National Diabetes Registry (CroDiab). The inclusion criteria were the following: diabetes diagnosis at < 45 years, age at sampling > 18 years, preserved endogenous insulin production (fasting C-peptide ≥ 0.2 nmol/L), and negative GAD antibodies. All the individuals from this cohort had *HNF1A* gene sequenced, identified allele variants systematically assessed [18] and aligned to the American College of Medical Genetics (ACMG) classification [24]. Patients with rare variants classified as pathogenic or likely pathogenic (ACMG classification 1–2) were categorized as HNF1A-MODY and analyzed within damaging (D) group ($N = 18$) and patients with likely benign or benign rare variants (ACMG classification 4–5) were analyzed within benign (B) group ($N = 9$). Specific variants subjects in these two groups carried are listed in Supplemental table 1. Other clinical characterization data is available in Supplemental table 2. This sample set was selected from a larger cohort based on sequencing results in order to include HNF1A mutations, along with controls without rare variants ($N = 71$). The study was approved by corresponding ethics committees and all participants gave written informed consent.

3. Materials and methods

3.1. Analysis of AGP N-glycosylation

Each sample underwent high-throughput and site-specific N-glycosylation analysis of AGP, as described previously [23]. The protocol was adapted for starting plasma volume of 20 μL instead of 50 μL. The only modifications - amounts of perchloric acid, RapiGest and TPCK-treated trypsin used in the protocol are listed below.

Starting 20 μL of plasma was mixed with 80 μL of 0.75 M perchloric acid. After centrifugation, supernatant containing the "seromucoid"

fraction, in which AGP is the dominant protein, was collected. By addition of phosphotungistic acid, "seromucoid" fraction was precipitated and supernatant discarded after centrifugation. The precipitate was further solubilized by addition of NaOH.

The solubilized precipitate was incubated with 2.5 μL of 1.5 % RapiGest to assist denaturation and later digestion. In order to break up disulfide bonds, cystine reduction with dithiothreitol and alkylation with iodoacetamide was performed. The pH was corrected by addition of ammonium bicarbonate after which the solution was incubated overnight with 0.8 μg of TPCK-treated trypsin to digest proteins. Afterwards the RapiGest was degraded by addition of HCl, and the obtained glycopeptides were enriched by binding to previously conditioned HILIC beads, washed and then eluted in ultrapure water with 0.1% TFA. Finally, eluted glycopeptides were dried down in a vacuum concentrator and kept at −20 °C until analysis.

Dried glycopeptides were reconstituted in ultrapure water and separated on a nanoACQUITY ultra-performance liquid chromatography (UPLC) instrument (Waters), coupled to Compact mass spectrometer (Bruker Daltonics). LaCyTools (version 1.0.11.0.1b.9) was used for automated quantification of acquired MS data, based on previously MS/MS confirmed glycopeptide structures. After the QC assessment, calculated absolute signal intensities were normalized to cumulative signal intensity per glycosylation site for each sample. In this step the inevitable variation of signal intensities between samples is removed from the final data and therefore sample comparison is made possible.

Considering five glycosylation sites present on the AGP molecule, along with two AGP isoforms present in human plasma, the abbreviations are used for naming individual AGP N-glycopeptides. The first Roman numeral presents corresponding glycosylation site, and the following Latin numeral presents contributing AGP isoform. The separated alphanumeric sequence describes the glycan composition by the number of different monosaccharide units present. The abbreviations are further explained in Supplemental Fig. 1.

3.2. Statistical analysis

To remove effect of variation in laboratory conditions during sample preparation and analysis, glycan data was batch corrected. Considering the non-normal distribution of the data, log-transformation was firstly performed. By applying the ComBat method (R package "sva"), the experimental noise was reduced, with sample plate defined within the model as the source of variation. The procedure was repeated for every glycopeptide, separately for each sample cohort.

To compare the levels of fucosylated AGP glycopeptides between different patient groups, generalized linear regression model was used. Log-transformed normalized glycopeptide data was set as dependent variable. The independent variables were type of diabetes for the diabetes type comparison cohort and *HNF1A* variant type along with sex for the second cohort. This was repeated for all analyzed fucosylated glycopeptides. Obtained effect size represents natural logarithm of difference in relative abundance of corresponding glycopeptide between selected patient groups. To control the false discovery rate, Benjamini-Hochberg method was applied, with adjusted p-value < 0.05 considered as significant.

Evaluation of discriminatory potential of fucosylated AGP glycopeptides for HNF1A-MODY and T2DM was performed using receiver operating characteristic curve (ROC) and the area under the curve (AUC) (R package "pROC"). Logistic regression models were created for each analyzed fucosylated glycopeptide with disease set as dependent and normalized glycopeptide data as an independent variable. The best performing glycopeptides were selected based on AUC calculated from ROC curves. To test the discriminatory performance of model including multiple fucosylated glycopeptides as independent variables, regularized (elastic net) regression model was used (R package "glmnet"). Before the training and validation of the model, regularization parameter alpha was tuned to 0.45 and lambda was selected based on minimal cross-validation prediction error rate. To evaluate the performance of the regularized logistic model, the 10-cross validation procedure was used, and AUC finally calculated. Comparison of AUCs between

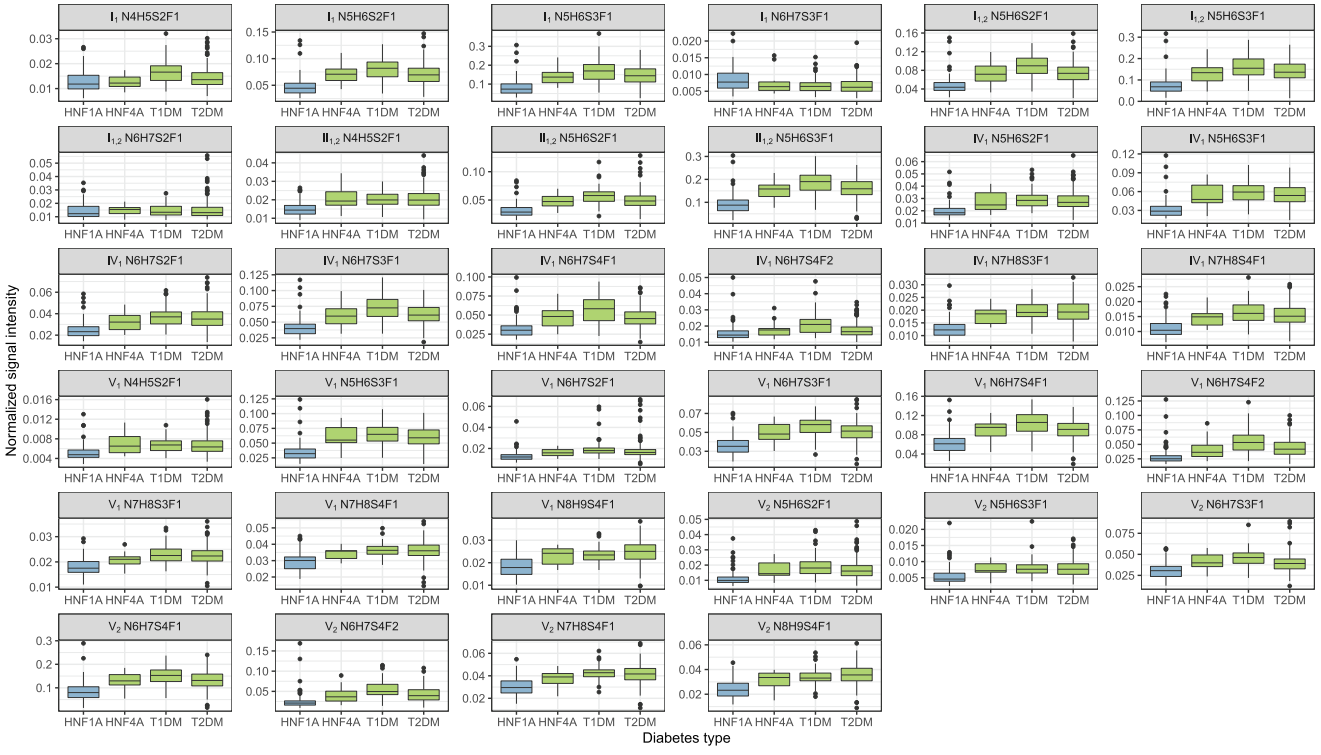

Fig. 1. Relative abundance of individual fucosylated AGP glycopeptides presented as boxplots and separated by diabetes type.

individual glycopeptide model and regularized logistic regression model was performed by bootstrap test with 2000 replicates (R package "pROC").

All data analysis and visualization were done using R programming language (version 3.6.3).

4. Results and discussion

Analysis of AGP N-glycosylation was performed on a glycopeptide level, thus providing site-specific information and allowing for a more comprehensive inspection.

Alpha-1-acid glycoprotein N-glycosylation was compared in respect to different diabetes subtypes as well as to predicted pathogenicity of *HNF1A* variants, considering presence and phenotypic effects of mutations. In addition, discriminatory potential of AGP N-glycopeptides in terms of identifying presence of damaging *HNF1A* mutations was assessed.

Considering the known association between *HNF1A* gene and fucosylation of plasma proteins as well as results of previous studies on plasma protein N-glycosylation, the most substantial effects were expected in abundance of fucosylated AGP N-glycopeptides. Therefore, only fucosylated glycopeptides were compared between patients, while glycopeptides not carrying fucose were excluded from the analysis.

4.1. Glycosylation comparison between various types of diabetes

First, we compared abundances of 34 fucosylated AGP glycopeptides between individuals with various diabetes types: T1DM, T2DM, HNF1A-MODY and HNF4A-MODY to test the HNF1A-MODY-associated decrease in AGP fucosylation, as hypothesized based on previous studies.

Type 2 diabetes is the most common misdiagnosis for HNF1A-MODY, making the comparison of these two diabetes subtypes the most clinically relevant. As visible from Fig. 1 and Table 1, the relative abundance of almost every compared fucosylated glycopeptide is reduced in individuals with HNF1A-MODY when compared to other types of diabetes. The lowest adjusted p-value of 2.34×10^{-44} was calculated for IV_1 N7H8S3F1 when compared to T2DM.

The effect size and very high statistical significance of the observed results are confirming a strong association between HNF1A-MODY and fucosylation levels, making AGP glycopeptide analysis a prospective diagnostic tool candidate for HNF1A-MODY. Furthermore, our data are in agreement with previously published results on released plasma protein N-glycans [17].

The observed differences in AGP glycopeptides are similarly pronounced in T1DM, indicating that this biomarker also provides adequate discrimination between HNF1A-MODY and T1DM. Therefore, this test could be useful in the 15–20% of adults with T1DM who do not have measurable islet-cell antibodies.

HNF4A and HNF1A as transcription factors regulate the expression of many overlapping genes [25] and HNF4A has also been identified to partially regulate plasma protein fucosylation [16]. Nevertheless, our results show there is difference in AGP fucosylation between HNF4A-MODY and HNF1A-MODY patients, statistically significant for almost every glycopeptide (Table 1). Sequencing panels currently used in most testing labs cover both HNF1A and HNF4A-MODY, therefore making the clinical value of this comparison less pronounced.

The observed effect appears to be consistent across all AGP glycosylation sites showing the extent to which *HNF1A* affects fucosylation. This protein-wide effect is also likely the reason for previously described clear visibility of the change in total plasma protein glycan analysis. The only glycan structure changing oppositely from the expected direction (i.e., increasing relative abundancy in HNF1A-MODY patients) is

Table 1
Difference in relative abundance of individual glycopeptides in different diabetes types compared to HNF1A-MODY. Results are shown as coefficients representing natural logarithm of difference in abundance and corresponding p-values.

Site	Glycan	Coefficient HNF4A	p-value HNF4A	p adjusted HNF4A	Coefficient T1DM	p-value T1DM	p adjusted T1DM	Coefficient T2DM	p-value T2DM	p adjusted T2DM
I_1	N4H5S2F1	0.021	7.79E-01	7.79E-01	**0.262**	**3.61E-08**	**3.84E-08**	**0.110**	**1.52E-03**	**1.56E-03**
I_1	N5H6S2F1	**0.404**	**2.07E-07**	**6.41E-07**	**0.553**	**7.82E-27**	**1.90E-26**	**0.400**	**5.70E-26**	**7.75E-26**
I_1	N5H6S3F1	**0.595**	**8.08E-08**	**3.43E-07**	**0.750**	**7.79E-25**	**1.56E-24**	**0.606**	**8.82E-29**	**1.58E-28**
I_1	N6H7S3F1	−0.113	2.27E-01	2.41E-01	**−0.170**	**3.57E-03**	**3.68E-03**	**−0.179**	**3.79E-05**	**4.03E-05**
$I_{1,2}$	N5H6S2F1	**0.425**	**5.98E-08**	**2.90E-07**	**0.630**	**1.19E-36**	**1.35E-35**	**0.448**	**1.48E-34**	**6.30E-34**
$I_{1,2}$	N5H6S3F1	**0.613**	**1.85E-08**	**1.26E-07**	**0.796**	**1.88E-31**	**1.07E-30**	**0.674**	**2.11E-39**	**1.43E-38**
$I_{1,2}$	N6H7S2F1	0.063	4.60E-01	4.74E-01	0.027	5.90E-01	5.90E-01	0.045	2.21E-01	2.21E-01
$II_{1,2}$	N4H5S2F1	**0.324**	**3.00E-07**	**8.50E-07**	**0.304**	**1.50E-15**	**1.76E-15**	**0.320**	**2.97E-28**	**4.80E-28**
$II_{1,2}$	N5H6S2F1	**0.446**	**6.03E-09**	**6.83E-08**	**0.618**	**5.62E-37**	**9.56E-36**	**0.472**	**3.66E-39**	**2.07E-38**
$II_{1,2}$	N5H6S3F1	**0.560**	**6.25E-10**	**1.06E-08**	**0.739**	**6.67E-38**	**2.27E-36**	**0.583**	**3.31E-42**	**3.75E-41**
IV_1	N5H6S2F1	**0.313**	**1.86E-06**	**3.51E-06**	**0.393**	**1.70E-22**	**2.89E-22**	**0.351**	**7.87E-31**	**1.91E-30**
IV_1	N5H6S3F1	**0.495**	**2.63E-08**	**1.49E-07**	**0.642**	**8.58E-31**	**3.72E-30**	**0.572**	**8.46E-42**	**7.19E-41**
IV_1	N6H7S2F1	**0.271**	**1.15E-04**	**1.45E-04**	**0.411**	**1.67E-21**	**2.71E-21**	**0.364**	**3.99E-29**	**7.98E-29**
IV_1	N6H7S3F1	**0.382**	**3.81E-07**	**9.97E-07**	**0.543**	**8.76E-31**	**3.72E-30**	**0.413**	**2.26E-32**	**6.99E-32**
IV_1	N6H7S4F1	**0.371**	**1.86E-06**	**3.51E-06**	**0.558**	**2.32E-30**	**8.78E-30**	**0.379**	**2.02E-26**	**2.86E-26**
IV_1	N6H7S4F2	**0.136**	**2.97E-02**	**3.26E-02**	**0.294**	**1.30E-14**	**1.47E-14**	**0.115**	**2.75E-05**	**3.02E-05**
IV_1	N7H8S3F1	**0.351**	**1.62E-08**	**1.26E-07**	**0.435**	**1.80E-29**	**5.56E-29**	**0.423**	**6.87E-46**	**2.34E-44**
IV_1	N7H8S4F1	**0.282**	**6.01E-06**	**9.73E-06**	**0.376**	**1.18E-22**	**2.11E-22**	**0.338**	**1.49E-31**	**4.22E-31**
V_1	N4H5S2F1	**0.289**	**1.50E-05**	**2.22E-05**	**0.285**	**1.02E-12**	**1.11E-12**	**0.269**	**3.41E-19**	**4.09E-19**
V_1	N5H6S3F1	**0.548**	**5.05E-10**	**1.06E-08**	**0.650**	**3.48E-32**	**2.36E-31**	**0.577**	**2.12E-43**	**3.60E-42**
V_1	N6H7S2F1	**0.274**	**5.72E-04**	**6.71E-04**	**0.439**	**1.97E-19**	**2.68E-19**	**0.318**	**7.36E-19**	**8.34E-19**
V_1	N6H7S3F1	**0.322**	**9.18E-08**	**3.47E-07**	**0.450**	**1.64E-32**	**1.40E-31**	**0.340**	**7.15E-34**	**2.70E-33**
V_1	N6H7S4F1	**0.383**	**1.85E-07**	**6.28E-07**	**0.521**	**6.39E-30**	**2.17E-29**	**0.389**	**1.54E-30**	**3.50E-30**
V_1	N6H7S4F2	**0.378**	**3.15E-05**	**4.29E-05**	**0.618**	**1.37E-27**	**3.57E-27**	**0.435**	**1.53E-25**	**2.00E-25**
V_1	N7H8S3F1	**0.138**	**1.18E-03**	**1.34E-03**	**0.242**	**2.91E-20**	**4.30E-20**	**0.220**	**7.28E-29**	**1.38E-28**
V_1	N7H8S4F1	**0.172**	**2.54E-05**	**3.60E-05**	**0.226**	**1.03E-19**	**1.46E-19**	**0.216**	**2.44E-30**	**5.18E-30**
V_1	N8H9S4F1	**0.263**	**8.70E-07**	**1.97E-06**	**0.267**	**1.34E-16**	**1.75E-16**	**0.324**	**6.42E-38**	**3.12E-37**
$V_{1,2}$	N5H6S2F1	**0.399**	**1.70E-06**	**3.51E-06**	**0.541**	**1.02E-25**	**2.32E-25**	**0.424**	**1.85E-28**	**3.15E-28**
$V_{1,2}$	N5H6S3F1	**0.400**	**4.35E-06**	**7.39E-06**	**0.430**	**3.24E-16**	**4.04E-16**	**0.411**	**4.58E-25**	**5.77E-25**
$V_{1,2}$	N6H7S3F1	**0.305**	**6.77E-06**	**1.05E-05**	**0.425**	**3.30E-24**	**6.23E-24**	**0.273**	**3.49E-19**	**4.09E-19**
$V_{1,2}$	N6H7S4F1	**0.471**	**6.91E-07**	**1.68E-06**	**0.609**	**2.48E-25**	**5.27E-25**	**0.472**	**1.84E-27**	**2.72E-27**
$V_{1,2}$	N6H7S4F2	**0.546**	**3.21E-06**	**5.75E-06**	**0.814**	**9.96E-29**	**2.82E-28**	**0.594**	**3.11E-28**	**4.80E-28**
$V_{1,2}$	N7H8S4F1	**0.245**	**4.09E-05**	**5.35E-05**	**0.349**	**1.97E-21**	**3.04E-21**	**0.324**	**2.25E-31**	**5.89E-31**
$V_{1,2}$	N8H9S4F1	**0.258**	**1.80E-04**	**2.18E-04**	**0.341**	**3.32E-16**	**4.04E-16**	**0.387**	**3.20E-33**	**1.09E-32**

I_1N6H7S3F1. This finding, although interesting, is challenging to interpret. However, it is the least abundant glycan structure characterized at this glycosylation site, so the reliability along with relevance of this observation might be limited.

Even though the relation of HNF1A-MODY to type 2 diabetes has the largest clinical significance, it was also important to confirm the observed shift in glycosylation is not present in other types of diabetes because this contributes to positive predictive value of testing. On the other hand, inflammatory diseases often result in increased levels of AGP fucosylation and such conditions should be considered in case of clinical use [26–28].

4.2. AGP glycosylation in young-onset non-autoimmune diabetes including both HNF1A-MODY and T2DM

Individuals with early-onset diabetes characterized by preserved insulin production and absence of pancreatic autoimmunity are potential HNF1A-MODY cases, even though many more with T2DM match these criteria. Since this similarity of traits is a major source of HNF1A-MODY misdiagnoses, it is important to find a way of identifying true HNF1A-MODY cases from a large pool of candidates. Therefore, AGP N-glycosylation analysis was performed on a group of 98 individuals matching the previous description, all of whom had *HNF1A* previously sequenced. It was also previously confirmed that 18 subjects carry damaging *HNF1A* mutations. The aim of this experiment was to test if the change in AGP fucosylation corresponds to functionality of *HNF1A* gene, i.e., if AGP fucosylation alterations could detect disease-causing *HNF1A* variants. In addition, assessment of AGP glycopeptides' ability to distinguish patients with HNF1A-MODY from a population of patients with non-autoimmune diabetes was performed by ROC curve analysis.

Levels of AGP's fucosylated glycopeptides were compared between individuals with pathogenic or likely pathogenic *HNF1A* variants and those with predicted normal function of the protein, separately for subjects with benign or likely benign variants and those without rare *HNF1A* variants.

Again, our results showed clear separation of patients carrying deleterious mutations, who exhibited lower levels of fucosylated AGP glycopeptides (Fig. 2). Linear regression confirmed the observed effect to be highly statistically significant for majority of glycopeptides and IV_1 N7H8S3F1 again showed the largest statistical significance (p-value 1.46×10^{-10} for comparison with non-rare variants) (Table 2). This structure was not reported in plasma N-glycosylation study previously performed on these subjects [18]. Even though highly branched and complex glycans indicated the changes in AGP glycosylation, by specific AGP glycopeptide analysis, we were able to observe structures not available by previously used method and thereby potentially find markers with even better diagnostic performance.

4.3. Single AGP glycopeptide as a predictor of HNF1A-MODY

To examine the classification performance of fucosylated AGP glycopeptides in identifying the patients with pathogenic or likely pathogenic *HNF1A* variants among patients with benign or without rare variants, ROC curve analysis was conducted. Individual glycopeptides were tested for their discriminative power and, as before, IV_1 N7H8S3F1 showed the best performance with AUC of 0.94 (95% CI 0.90–0.99) (Fig. 3 A). Classification performance of the remaining glycopeptides is shown in Supplemental table 3. The previously described and suggested HNF1A-MODY diagnostic tool based on total plasma protein N-glycans [18] requires relatively complex procedure involving N-glycan release, labeling, and clean-up, followed by UPLC analysis and data processing, because many analytes contribute to the final result. Also, the best performing plasma glycan peak had an AUC of 0.90, lower than single AGP N-glycopeptide. Previous comparison of other MODY-related biomarkers including hsCRP, 1,5-anhydroglucitol, cystatin C, HDL and several others, found hsCRP to be the most promising [29]. However both in study performed on total plasma N-glycosylation [18] and in this study, glycan based biomarkers outperformed hsCRP (AUC: 0.83,

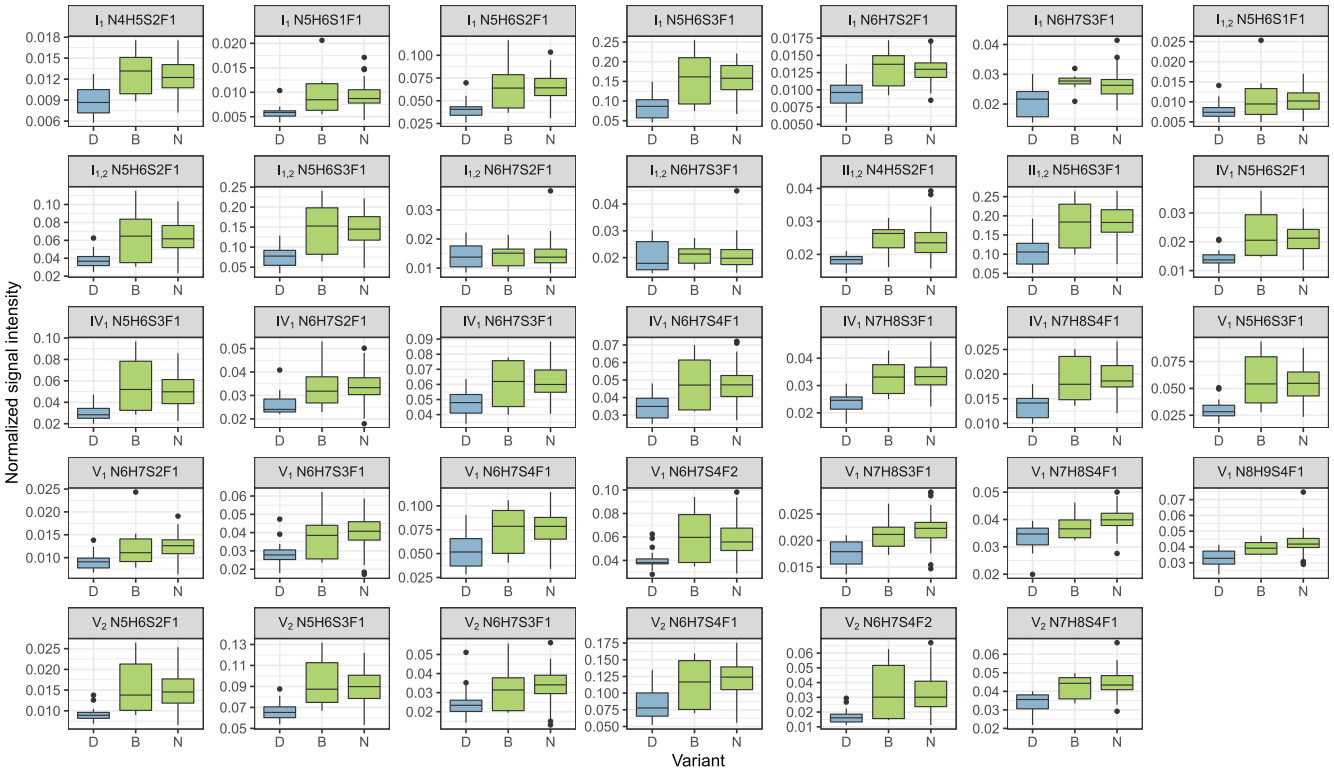

Fig. 2. Relative abundance of individual fucosylated AGP glycopeptides presented as boxplots and separated based on the degree of HNF1A variant pathogenicity: pathogenic or likely pathogenic (D), benign or likely benign (B) and subjects without rare variants (N).

Table 2
Difference in relative abundance of individual glycopeptides in patients with benign *HNF1A* variants (B) and without any allele variants(N) compared to damaging *HNF1A* variants. Results are shown as coefficients representing natural logarithm of difference in abundance and corresponding p-values.

Site	Glycan	Coefficient B	p-value B	p adjusted B	Coefficient N	p-value N	p adjusted N
I_1	N4H5S2F1	**0.241**	**9.44E-03**	**1.69E-02**	**0.302**	**5.59E-07**	**8.64E-07**
I_1	N5H6S1F1	**0.367**	**4.50E-03**	**1.18E-02**	**0.421**	**4.72E-07**	**8.03E-07**
I_1	N5H6S2F1	**0.328**	**5.45E-03**	**1.30E-02**	**0.447**	**1.21E-08**	**4.10E-08**
I_1	N5H6S3F1	**0.408**	**6.14E-03**	**1.30E-02**	**0.569**	**9.36E-09**	**3.54E-08**
I_1	N6H7S2F1	**0.265**	**1.76E-03**	**6.65E-03**	**0.315**	**1.54E-08**	**4.75E-08**
I_1	N6H7S3F1	**0.241**	**8.60E-03**	**1.69E-02**	**0.234**	**5.30E-05**	**6.01E-05**
$I_{1,2}$	N5H6S1F1	0.164	1.89E-01	2.01E-01	**0.278**	**6.35E-04**	**6.74E-04**
$I_{1,2}$	N5H6S2F1	**0.294**	**2.46E-02**	**3.81E-02**	**0.469**	**1.08E-07**	**2.30E-07**
$I_{1,2}$	N5H6S3F1	**0.410**	**4.13E-03**	**1.17E-02**	**0.578**	**3.16E-09**	**1.53E-08**
$I_{1,2}$	N6H7S2F1	0.009	9.37E-01	9.37E-01	0.030	6.86E-01	6.86E-01
$I_{1,2}$	N6H7S3F1	0.033	7.15E-01	7.37E-01	0.026	6.42E-01	6.61E-01
$II_{1,2}$	N4H5S2F1	**0.253**	**9.59E-04**	**6.63E-03**	**0.251**	**6.27E-07**	**9.28E-07**
$II_{1,2}$	N5H6S3F1	**0.406**	**1.21E-03**	**6.63E-03**	**0.552**	**1.80E-10**	**3.06E-09**
IV_1	N5H6S2F1	**0.318**	**1.48E-03**	**6.63E-03**	**0.368**	**4.06E-08**	**9.85E-08**
IV_1	N5H6S3F1	**0.407**	**1.56E-03**	**6.63E-03**	**0.486**	**1.90E-08**	**5.37E-08**
IV_1	N6H7S2F1	0.157	5.55E-02	6.53E-02	**0.241**	**9.30E-06**	**1.13E-05**
IV_1	N6H7S3F1	0.148	5.57E-02	6.53E-02	**0.252**	**1.10E-06**	**1.50E-06**
IV_1	N6H7S4F1	**0.210**	**2.22E-02**	**3.59E-02**	**0.301**	**8.28E-07**	**1.17E-06**
IV_1	N7H8S3F1	**0.260**	**1.18E-04**	**4.02E-03**	**0.326**	**4.30E-12**	**1.46E-10**
IV_1	N7H8S4F1	**0.258**	**1.30E-03**	**6.63E-03**	**0.335**	**9.49E-10**	**6.45E-09**
V_1	N5H6S3F1	**0.483**	**3.06E-04**	**5.20E-03**	**0.574**	**3.32E-10**	**3.10E-09**
V_1	N6H7S2F1	0.191	4.46E-02	5.83E-02	**0.272**	**1.49E-05**	**1.74E-05**
V_1	N6H7S3F1	0.198	6.22E-02	7.05E-02	**0.341**	**1.59E-06**	**2.08E-06**
V_1	N6H7S4F1	0.238	3.77E-02	5.34E-02	**0.385**	**5.37E-07**	**8.64E-07**
V_1	N6H7S4F2	**0.230**	**3.24E-02**	**4.79E-02**	**0.318**	**7.66E-06**	**9.64E-06**
V_1	N7H8S3F1	**0.146**	**8.98E-03**	**1.69E-02**	**0.221**	**6.67E-09**	**2.83E-08**
V_1	N7H8S4F1	0.103	4.08E-02	5.54E-02	**0.183**	**7.02E-08**	**1.59E-07**
V_1	N8H9S4F1	**0.178**	**5.91E-03**	**1.30E-02**	**0.269**	**1.50E-09**	**8.52E-09**
$V_{1,2}$	N5H6S2F1	**0.384**	**1.41E-03**	**6.63E-03**	**0.447**	**3.00E-08**	**7.85E-08**
$V_{1,2}$	N5H6S3F1	**0.235**	**1.99E-03**	**6.77E-03**	**0.267**	**1.34E-07**	**2.68E-07**
$V_{1,2}$	N6H7S3F1	0.163	1.80E-01	1.98E-01	**0.321**	**6.30E-05**	**6.90E-05**
$V_{1,2}$	N6H7S4F1	0.216	4.74E-02	5.97E-02	**0.376**	**2.93E-07**	**5.25E-07**
$V_{1,2}$	N6H7S4F2	**0.382**	**1.87E-02**	**3.19E-02**	**0.566**	**2.29E-07**	**4.32E-07**
$V_{1,2}$	N7H8S4F1	**0.183**	**2.71E-03**	**8.37E-03**	**0.264**	**3.64E-10**	**3.10E-09**

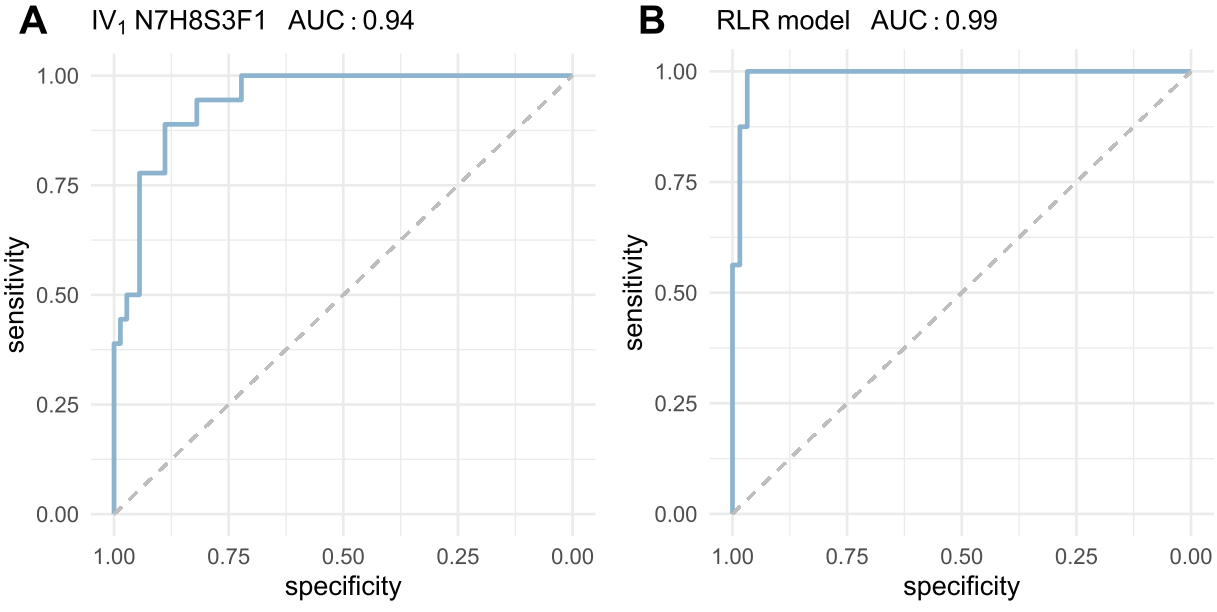

Fig. 3. Classification performance of HNF1A-MODY prediction models based on the best performing individual glycopeptide (A) and on multiple glycopeptides (B) selected by regularized logistic regression (RLR), visualized by ROC curves and corresponding areas under the curve (AUC).

Supplemental table 3.) apart from hsCRP being routinely available. The promising performance of IV_1 N7H8S3F1 introduces a potential for development of a diagnostic tool based on a single glycopeptide, without the need to analyze many glycopeptides or glycans as performed in this and previous studies. Single glycopeptide approach would require less sample preparation and less complex analytical system, while still retaining the benefits of glycopeptide analysis compared to released glycans. This would make the analysis much simpler and more achievable for clinical application.

4.4. Multiple AGP glycopeptides as predictors of HNF1A-MODY

Next, the predictive model based on glycopeptides selected using the regularized logistic regression from the complete AGP glycopeptide profile was tested. Glycopeptides included in the model along with corresponding coefficients are listed in the Supplemental table 3. The obtained ROC curve had AUC of 0.99 (95% CI 0.98–1), suggesting even better performance than IV_1 N7H8S3F1 glycopeptide but not reaching statistical significance (p-value = 0.056) when compared (Fig. 3 B). This shows that although individual glycopeptides are highly informative, additional information is likely contained within the total glycosylation profile of APG. Nevertheless, analysis and comparison of total N-glycosylation of AGP requires more elaborate sample preparation and data analysis, possibly not being justified for clinical environment.

The observed difference in glycopeptide levels between functional and disease-causing *HNF1A* allele variants, represents an advantage since *HNF1A* sequencing results require further interpretation, sometimes highly demanding for previously uncharacterized variants. Moreover, it could help to interpret the phenotypical effect for variants of unknown significance.

4.5. Where could measurement of AGP N-glycosylation fit into a diagnostic pathway for MODY?

We have discussed above the potential role of the AGP N-glycosylation test to assist with variant assessment after sequencing. In addition, the test could also be used for screening individuals not currently eligible for diagnostic testing to identify those at higher likelihood of carrying a functional *HNF1A* variant. Most testing labs currently use next-generation gene sequencing panel, typically including up to 20–30 genes associated with diabetes alone or with diabetes syndromes. The eligibility criteria for such testing commonly include younger age of onset (up to age 25–35), with parental history of diabetes. Those considered by clinicians more likely to have type 2 diabetes e.g. with older age of onset, obese individuals or those from non-white ethnic groups may be either excluded or less often offered testing for monogenic diabetes. For this reason, inclusion of AGP fucosylation data into currently used screening tools such as MODY probability calculator [30] could be beneficial.

5. Conclusions

The AGP N-glycosylation analysis in individuals affected by HNF1A-MODY, in comparison to other diabetes subtypes, confirmed the expected significant reduction in fucosylation. Furthermore, the association of fucosylated glycopeptide levels with rare allele variants was demonstrated, where patients with disease-causing variants had decreased levels of fucosylated glycopeptides, compared to patients with benign, likely benign variants or without any. This represents a basis for proposal of AGP glycopeptide analysis as a diagnostic tool for HNF1A-MODY. The ROC curve analysis demonstrated a very good performance of individual glycopeptides, especially IV_1 N7H8S3F1, showing that single glycopeptide might serve as diagnostic and/or stratification tool.

Introduction of IV_1 N7H8S3F1 as a single glycopeptide marker would allow for effective identification of patients with dysfunctional *HNF1A* and therefore offer a path to definitive diagnosis, regardless of potentially inconclusive sequencing results. Also, it could be used for patients currently excluded from sequencing considering the likely low price of analysis (for method used in this study, excluding MS instrumentation, cost per sample was under 5 euro). Compared to the previously proposed total plasma protein N-glycosylation analysis, single glycopeptide-based test carries an advantage in terms of increased signal specificity. Furthermore, being a methodologically simpler approach, it would more likely be able to take advantage of MS-based systems' increasing presence in clinics and be a cost-effective addition to the whole differential diagnosis process.

This way patients less likely carrying damaging *HNF1A* mutations and those currently excluded from guidelines for *HNF1A* sequencing (e.g. older age at detection of glycemic dysregulation), could routinely be tested and otherwise missed HNF1A-MODY cases still diagnosed.

Author contributions

OG conceived and supervised the study. OG, KRO and GL designed the study. MT and TŠ acquired the data. MT, TŠ, MN, TK and OG analyzed and interpreted data. MT, OG and KRO wrote the manuscript. PRN, ET, JS, DG, EPM, AJ and KRO enrolled patients. AJ and KRO stratified patients. AJ sequenced *HNF1A* from 98 participants from UK and Croatia. All authors reviewed the manuscript and approved the final version of the manuscript. OGK is the guarantor of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Funding

This work was supported by the European Structural and Investment funding for the 'Croatian National Centre of Research Excellence in Personalized Healthcare' [grant number KK.01.1.1.01.0010], European Foundation for Study of Diabetes/Novo Nordisk Research Grant and the European Regional Development Fund grant 'CardioMetabolic' agreement [grant number KK.01.2.1.02.0321], VEGA [grant number 1/0211/18] and Slovak Ministry of Health [grant number 2019/20- LFUK-8]. The funding sources had no involvement in study design.

Declaration of Competing Interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: G. L. is the founder and owner and M.N. and T.Š are employees of Genos Ltd, a company that specializes in high-throughput glycomics and has several patents in this field. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.diabres.2022.109226.

References:

[1] Flannick J, Johansson S, Njølstad PR. Common and rare forms of diabetes mellitus: towards a continuum of diabetes subtypes. Nat Rev Endocrinol 2016;12(7):394–406.
[2] Anik A, Çatli G, Abaci A, and Böber E. "Maturity-onset diabetes of the young (MODY): An update," *Journal of Pediatric Endocrinology and Metabolism*, vol. 28, no. 3–4. Walter de Gruyter GmbH, pp. 251–263, Mar. 01, 2015, doi: 10.1515/jpem-2014-0384.
[3] Urakami T. "Maturity-onset diabetes of the young (MODY): Current perspectives on diagnosis and treatment". In: Diabetes, Metabolic Syndrome and Obesity: Targets and Therapy, vol. 12. Dove Medical Press Ltd.; 2019. p. 1047–56. https://doi.org/10.2147/DMSO.S179793.
[4] Shields BM, Hicks S, Shepherd MH, Colclough K, Hattersley AT, Ellard S. Maturity-onset diabetes of the young (MODY): How many cases are we missing? Diabetologia Dec. 2010;53(12):2504–8. https://doi.org/10.1007/s00125-010-1799-4.
[5] Kavvoura FK and Owen KR. "Maturity onset diabetes of the young: clinical characteristics, diagnosis and management.," *Pediatric endocrinology reviews : PER*, vol. 10, no. 2. pp. 234–242, Dec. 01, 2012.
[6] Pearson ER, Starkey BJ, Powell RJ, Gribble FM, Clark PM, Hattersley AT. Genetic cause of hyperglycaemia and response to treatment in diabetes. Lancet Oct. 2003;362(9392):1275–81. https://doi.org/10.1016/S0140-6736(03)14571-0.
[7] Colclough K, Bellanne-Chantelot C, Saint-Martin C, Flanagan SE, Ellard S. Mutations in the Genes Encoding the Transcription Factors Hepatocyte Nuclear Factor 1 Alpha and 4 Alpha in Maturity-Onset Diabetes of the Young and Hyperinsulinemic Hypoglycemia. Hum. Mutat. May 2013;34(5):669–85. https://doi.org/10.1002/humu.22279.

[8] Harries LW. *et al.*, "Isomers of the TCF1 gene encoding hepatocyte nuclear factor-1 alpha show differential expression in the pancreas and define the relationship between mutation position and clinical phenotype in monogenic diabetes," *Hum. Mol. Genet.*, vol. 15, no. 14, pp. 2216–2224, Jul. 2006, doi: 10.1093/hmg/ddl147.
[9] Flannick J, Beer NL, Bick AG, Agarwala V, Molnes J, Gupta N, et al. Assessing the phenotypic effects in the general population of rare variants in genes for a dominant Mendelian form of diabetes. Nat Genet 2013;45(11):1380–5.
[10] Varki A, et al. Essentials of Glycobiology. Cold Spring Harbor Laboratory Press; 2017.
[11] Keser T, Gornik I, Vučković F, Selak N, Pavić T, Lukić E, et al. Increased plasma N-glycome complexity is associated with higher risk of type 2 diabetes. Diabetologia 2017;60(12):2352–60. https://doi.org/10.1007/s00125-017-4426-9.
[12] Dotz V, Lemmers RFH, Reiding KR, Hipgrave Ederveen AL, Lieverse AG, Mulder MT, et al. Plasma protein N-glycan signatures of type 2 diabetes. Biochim. Biophys. Acta - Gen. Subj. 2018;1862(12):2613–22. https://doi.org/10.1016/j.bbagen.2018.08.005.
[13] Wittenbecher C, Štambuk T, Kuxhaus O, Rudman N, Vučković F, Štambuk J, et al. Plasma N-glycans as emerging biomarkers of cardiometabolic risk: A prospective investigation intheepic-potsdamcohortstudy. Diabetes Care 2020;43(3):661–8. https://doi.org/10.2337/dc19-1507.
[14] Pavić T. *et al.*, "Maturity onset diabetes of the young due to HNF1A variants in Croatia," *Biochem. Medica*, vol. 28, no. 2 Special Issue, 2018, doi: 10.11613/BM.2018.020703.
[15] Rudman N, Gornik O, and Lauc G. "Altered N-glycosylation profiles as potential biomarkers and drug targets in diabetes," *FEBS Letters*, vol. 593, no. 13. Wiley Blackwell, pp. 1598–1615, Jul. 01, 2019, doi: 10.1002/1873-3468.13495.
[16] Lauc G. *et al.*, "Genomics meets glycomics-the first gwas study of human N-glycome identifies HNF1A as a master regulator of plasma protein fucosylation," *PLoS Genet.*, vol. 6, no. 12, pp. 1–14, 2010, doi: 10.1371/journal.pgen.1001256.
[17] Thanabalasingham G, Huffman JE, Kattla JJ, Novokmet M, Rudan I, Gloyn AL, et al. Mutations in HNF1A result in marked alterations of plasma glycan profile. Diabetes 2013;62(4):1329–37. https://doi.org/10.2337/db12-0880.
[18] Juszczak A, Pavić T, Vučković F, Bennett AJ, Shah N, Pape Medvidović E, et al. Plasma fucosylated glycans and C-reactive protein as biomarkers of HNF1A-MODY in young adult–onset nonautoimmune diabetes. Diabetes Care 2019;42(1):17–26. https://doi.org/10.2337/dc18-0422.
[19] Clerc F, Reiding KR, Jansen BC, Kammeijer GSM, Bondt A, Wuhrer M. Human plasma protein N-glycosylation. Glycoconj. J. Jun. 2016;33(3):309–43. https://doi.org/10.1007/s10719-015-9626-2.
[20] Schmid K, Nimberg RB, Kimura A, Yamaguchi H, Binette JP. The carbohydrate units of human plasma α1-Acid glycoprotein. BBA - Protein Struct. Jun. 1977;492(2):291–302. https://doi.org/10.1016/0005-2795(77)90080-0.
[21] Schmid K. Human plasma alpha 1-acid glycoprotein–biochemical properties, the amino acid sequence and the structure of the carbohydrate moiety, variants and polymorphism. Prog Clin Biol Res. 1989;300:7–22.
[22] Dente L, Pizza MG, Metspalu A, Cortese R. Structure and expression of the genes coding for human alpha 1-acid glycoprotein. EMBO J. Aug. 1987;6(8):2289–96.
[23] Keser T, Tijardović M, Gornik I, Lukić E, Lauc G, Gornik O, et al. High-Throughput and Site-Specific N-Glycosylation Analysis of Human Alpha-1-Acid Glycoprotein Offers a Great Potential for New Biomarker Discovery. Mol. Cell. Proteomics 2021;20:100044. https://doi.org/10.1074/mcp.RA120.002433.
[24] Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and Guidelines for the Interpretation of Sequence Variants: A Joint Consensus Recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet. Med. 2015;17(5):405–24.
[25] Eeckhoute J, Formstecher P, Laine B. Hepatocyte nuclear factor 4α enhances the hepatocyte nuclear factor 1α-mediated activation of transcription. Nucleic Acids Res. 2004;32(8):2586–93. https://doi.org/10.1093/nar/gkh581.
[26] Hashimoto S, Asao T, Takahashi J, Yagihashi Y, Nishimura T, Saniabadi AR, et al. alpha1-acid glycoprotein fucosylation as a marker of carcinoma progression and prognosis. Cancer 2004;101(12):2825–36.
[27] Poland D, Schalkwijk C, Stehouwer C, Koeleman C, van het Hof B, van Dijk W. Increased alpha3-fucosylation of alpha1-acid glycoprotein in Type I diabetic patients is related to vascular function. Glycoconj. J. 2001;18(3):261–8. https://doi.org/10.1023/A:1012412908983.
[28] Rydén I, Påhlsson P, Lundblad A, Skogh T. Fucosylation of α1-acid glycoprotein (orosomucoid) compared with traditional biochemical markers of inflammation in recent onset rheumatoid arthritis. Clin. Chim. Acta Mar. 2002;317(1–2):221–9. https://doi.org/10.1016/S0009-8981(01)00803-8.
[29] Mughal SA, Thanabalasingham G, Owen KR. Biomarkers currently used for the diagnosis of maturity-onset diabetes of the young. Diabetes Management 2013;3(1):71–80.
[30] Shields BM, McDonald TJ, Ellard S, Campbell MJ, Hyde C, Hattersley AT. The development and validation of a clinical prediction model to determine the probability of MODY in patients with young-onset diabetes. Diabetologia May 2012;55(5):1265–72. https://doi.org/10.1007/S00125-011-2418-8.