Jin *et al. Clinical Proteomics* (2024) 21:68
https://doi.org/10.1186/s12014-024-09522-4

Clinical Proteomics

RESEARCH Open Access

Quantitative site-specific N-glycosylation analysis reveals IgG glyco-signatures for pancreatic cancer diagnosis

Yi Jin[3†], Ran Hu[3†], Yufan Gu[1], Ailin Wei[2], Ang Li[2,3*] and Yong Zhang[1*]

Abstract

Background Pancreatic cancer is a highly aggressive tumor with a poor prognosis due to a low early detection rate and a lack of biomarkers. Most of pancreatic cancer is pancreatic ductal adenocarcinoma (PDAC). Alterations in the N-glycosylation of plasma immunoglobulin G (IgG) have been shown to be closely associated with the onset and development of several cancers and could be used as biomarkers for diagnosis. The study aimed to explore intact N-glycosylation profile of IgG in patients with PDAC and find relation between intact N-glycosylation profile of IgG and clinical information such as diagnosis and prognosis.

Methods In this study, we employed a well-evaluated approach (termed GlycoQuant) to assess the site-specific N-glycosylation profile of human plasma IgG in both healthy individuals and patients with pancreatic ductal adenocarcinoma (PDAC). The datasets generated and analyzed during the current study are available in the ProteomeXchange Consortium (http://www.proteomexchange.org/) via the iProX partner repository, with the dataset identifier PXD051436.

Results The analysis of rapidly purified IgG samples from 100 patients with different stages of PDAC, in addition to 30 healthy controls, revealed that the combination of carbohydrate antigen 19−9 (CA19-9), IgG1-GP05 (IgG1-TKPREEQYNSTYR-HexNAc [4]Hex [5]Fuc [1]NeuAc [1]), and IgG4-GP04 (IgG4-EEQFNSTYR- HexNAc [4]Hex [5]Fuc [1] NeuAc [1]) can be used to distinguish between PDAC patients and healthy individuals (AUC = 0.988). In addition, cross validation of the diagnosis model showed satisfactory result.

Conclusions The study demonstrated that the integrated quantitative method can be utilized for large-scale clinical N-glycosylation research to identify novel N-glycosylated biomarkers. This could facilitate the development of clinical glycoproteomics.

Keywords Pancreatic ductal adenocarcinoma, Immunoglobulin G, N-glycosylation, Mass spectrometry

†Yi Jin and Ran Hua contributed equally to this study and were listed as co-first authors.

*Correspondence:
Ang Li
angli@scu.edu.cn
Yong Zhang
nankai1989@foxmail.com

[1]Department of Pancreatic Surgery and Institutes for Systems Genetics, West China Hospital, Sichuan University, Keyuan 4th Road, Gaopeng Avenue, Hi-tech Zone, Chengdu, Sichuan 610041, China
[2]Guang'an People's Hospital, Guang'an 638001, China
[3]Department of Pancreatic Surgery and Institutes for Systems Genetics, West China Hospital, Sichuan University, 37 Guoxue Alley, Chengdu, Sichuan 610041, China

Introduction

Pancreatic cancer is a highly malignant solid tumor with an overall five-year survival rate of less than 10% [1]. Additionally, the five-year survival rate for patients with metastasis is less than 3% [2]. It is important to note that more than 90% of pancreatic cancers are pancreatic ductal adenocarcinoma (PDAC) [3], and surgical resection is the only way to completely cure PDAC [4]. However, PDAC is often asymptomatic, and early symptoms are not typical. As a result, most patients are already in the advanced or distant metastatic stage at the time of diagnosis, missing the best surgical opportunity, or metastasizing and dying within one year of surgery [5]. Early diagnosis and timely treatment can effectively improve the surgical resection rate of PDAC and, consequently, improve the five-year survival rate of patients.

Changes to the structure or quantity of proteins can directly impact normal physiological activities and even lead to neoplastic lesions. Glycosylation is a crucial and intricate post-translational modification, with over half of all human proteins requiring glycosylation [6]. Glycosylation directly involves in almost all human diseases, including inflammation, infectious diseases, cancer, diabetes, and neurodegeneration [7–10]. In malignant tumors, some glycoproteins have abnormal glycosylation patterns and are involved in various processes of tumorigenesis, such as protein folding and degradation, angiogenesis, differentiation, cell growth, tumor cell dissociation, invasion, epithelial-mesenchymal transformation and metastasis [11]. Studies have shown that the response of glycosylation modification to disease may be earlier and more obvious compared to changes in protein expression [12]. Additionally, altered glycoproteins in cancer cells could be released into the blood, and consequently, easy to be detected. It is important to note that these findings are objective and not based on subjective evaluations. Therefore, the glycosylation related to specific tumor may become specific cancer biomarker [13]. The use of advanced glycoproteomics technology to identify characteristic changes in this process may provide key information for disease diagnosis, early detection, treatment, and prognosis assessment.

Traditionally, tumor biomarkers refer to molecules (such as nucleic acids, proteins, and metabolites) that are produced either directly by malignant cells or by other cells in response to certain malignant or non-malignant conditions. Currently, the biomarker commonly used for the diagnosis, therapeutic efficacy evaluation, benign and malignant screening, and prognostic evaluation of PDAC is serum carbohydrate antigen 19−9 (CA199) [14]. However, the specificity and sensitivity of CA19-9 in diagnosing PDAC were only 78.9% and 69.9%, respectively [15]. What's more, the positive expression of CA19-9 is only found in the later stages of tumor progression, making it difficult to detect PDAC early. Furthermore, CA19-9 levels may also increase in the presence of jaundice, pancreatitis, bile duct obstruction, and other malignant tumors such as colon, stomach, and lung cancer, leading to insufficient specificity [16–19]. Additionally, the expression of CA19-9 requires the presence of Lewis antigen. And in about 5–10% of patients, the lack of Lewis antigen results in the phenomenon of false negative [20]. Therefore, the exploration of new tumor biomarkers for PDAC in order to achieve more accurate and early diagnosis remains the focus and difficulty of current clinical research.

Plasma immunoglobulin G (IgG) is the most abundant immunoglobulin in the circulatory system. It plays an important role in antigen neutralization, phagocytosis, cytotoxicity, and maintenance of immune homeostasis. The pattern and level of N-glycosylation in the IgG Fc fragment are crucial for most of these IgG-mediated protective immune responses [21]. Changes in N-glycosylation are associated with various diseases, including autoimmune diseases, infections, and cancer [22]. N-glycosylation of IgG has attracted attention as a potential biomarker for the diagnosis of various diseases such as prostate carcinoma, type 2 diabetes and chronic kidney diseases [22–24]. However, to our knowledge, it has not been studied systematacially in PDAC. Thus, it is worthwhile to study the IgG glycosylation modification in PDAC and we hypothesize that N-glycosylation of IgG could have the potential to assist CA19-9 in PDAC diagnosis.

In this study, we used a well-evaluated approach (termed GlycoQuant), which integrated fast IgG preparation method, combined electron-transfer/higher-energy collisional dissociation (EThcD) and stepped collision energy/higher-energy collisional dissociation (sceHCD) dissociation method EThcD-sceHCD and updated PANDA software tool to enable the quantification of intact N-glycopeptides [25]. GlycoQuant is a set of processes designed by our research team [25] to complete intact N-glycopeptide quantification, including sample preparation, LC-MS/MS analysis, identification, quantification, and bioinformatics analysis. Moreover, in previous study, we confirmed that ETHCD-sceHCD can quantify more intact N-glycopeptides than EThcD or sceHCD alone, and the average proportion of missing values is lowest, which has strong applicability in clinical N-glycoproteomics studies. Based on GlycoQuant, the site-specific N-glycosylation modifications of human plasma IgG in healthy people and patients with PDAC were characterized. When combined with some clinical indicators, the differentially expressed intact N-glycopeptides of IgG may be utilized as accurate diagnostic tools.

Materials and methods

Biospecimen collection

The study was conducted in adherence to the Declaration of Helsinki and approved by the medical ethics committee at the West China Hospital, Sichuan University. Plasma samples were collected from healthy controls (HC=30) and patients with pancreatic cancer (PDAC=100) and stored at -80 ℃ before use. Written informed consent was obtained from all participants.

This prospective observational study collected plasma samples from patients admitted to the Department of Pancreatic Surgery of West China Hospital between June 2020 and May 2022. The inclusion criteria were as follows: (1) be over 18 years of age, (2) have accurate Chinese character writing ability, (3) be proficient in oral expression, (4) have no history of mental illness, (5) understand the purpose and content of the study, (6) be able to sign informed consent, and (7) have been diagnosed with PDAC for the first time in the Department of Pancreatic Surgery at West China Hospital of Sichuan University. The diagnostic gold standard was based on pathological diagnosis of intraoperative tissue samples. The exclusion criteria were as follows: (1) those with non-primary pancreatic tumors or with two or more types of tumors including pancreatic cancer, (2) those with a history of tumor or chemoradiotherapy, (3) those who need to take immunosuppressants or hormones for an extended period due to disease, and (4) those diagnosed with kidney disease such as nephrotic syndrome, IgA nephropathy, acute or chronic renal insufficiency, etc. (5) other pancreatic borderline or benign tumors include intraductal papillary mucinous neoplasm, pancreatic serous or mucinous cystadenoma, solid pseudopapillary tumor of the pancreas, and pancreatic neuroendocrine tumor.

Plasma samples were collected from patients diagnosed with PDAC prior to surgery. Mass spectrometry was performed after postoperative pathological diagnosis confirmed PDAC, with the consent of patients and their families.

Clinical data collection

The clinical data collected from patients included their age, sex, body mass index (BMI), hypertension, diabetes, smoking and drinking habits, jaundice, CA19-9 levels, tumor size, tumor location (head of pancreas, uncinate process of pancreas, neck of pancreas, body of pancreas, and tail of pancreas), and stage of PDAC. The classification and staging of PDAC were performed according to the (tumor–node–metastasis) TNM staging and the American joint committee on cancer (AJCC) staging (stage I, II, III, IV), edition 8 [26].

Preparation of IgG

Plasma IgG from 130 samples were purified using our optimized method previously described. In brief, we mixed 20 μL of plasma in a microcentrifuge tube with 40 μL of equilibrated immobilized protein A/G agarose. The mixture was then incubated with 200 μL of binding buffer (25 mM tristate buffer, 150 mM sodium chloride, pH7.2) for 2 h at 4 ℃ on a rotator. After a brief centrifugation, the supernatant was discarded, and we added approximately 500 μL of binding buffer to wash the unbound proteins. Subsequently, 50 μL of elution buffer (0.1 M formic acid) was added to release the captured plasma IgGs. The tube was incubated for 5 min at 25 ℃ on a rotator, and then briefly centrifuged to collect purified IgGs. Quantification was performed using the bicinchoninic acid protein assay at 562 nm.

Denature, reduction, alkylation and digestion

The 20 μg of IgGs was added to 200 μL of 50 mM ammonium bicarbonate (ABC) buffer (pH8.5), and the mixture was denatured by heating at 95 ℃ for 10 min. The protein was then reduced and alkylated by adding 4 μL of 1 M dithiothreitol for 45 min at 56 ℃, followed by 10 μL of 1 M iodoacetamide (IAM) for 30 min at 25 ℃ in the dark. The resulting mixture was then loaded into a 10 kDa ultrafiltration tube (Millipore Corporation, Billerica). The agents were subsequently replaced with ABC buffer three times and centrifuged at 13,000 g for 10 min. Next, approximately 500 ng of trypsin and 100 μL of ABC were added and oscillated overnight at 37 ℃. The digests were collected by centrifugation at 13,000 g for 10 min at 4 ℃. Finally, the peptides were dried using the SpeedVac and stored at -80 ℃ for further analysis.

LC-MS/MS analysis

The peptides from each sample were resuspended in 40 μL of 0.1% formic acid. Subsequently, 1 μL of peptides was separated on a ReproSil-Pur C18-AQ column (1.9 μm, 100 μm inner diameter, length 25 cm; Dr Maisch) using a 30-min gradient (0–2 min, 5–12% B; 2–7 min, 12–22% B; 7–21 min, 22–32% B; 21–22 min, 32–90% B; 22–30 min, 90% B) at a flow rate of 350 nL/min. Finally, the separated peptides were analyzed using an Orbitrap Fusion Lumos Mass Spectrometer (Thermo Fisher, USA).

An alternative fragmentation between the EThcD and sceHCD modes was used in a duty cycle of 3 s for EThcD-sceHCD-MS/MS. During the EThcD duty cycle of 2 s, MS1 was analyzed at an Orbitrap resolution of 60,000 within the 400–1600 m/z range. The RF lens, AGC target, MIT, and exclusion duration were set to 40%, 2.0 e5, 50 ms, and 15 s, respectively. MS2 was analyzed at an Orbitrap resolution of 30,000 with 2 m/z. The AGC target, MIT, and EThcD type were standard, 150 ms, and

35%, respectively. During the sceHCD duty cycle of 1 s, MS1 was analyzed using a range of 400–1600 m/z at an Orbitrap resolution of 60,000. The RF lens, AGC target, MIT, and exclusion duration were set to 40%, standard, auto, and 15 s, respectively. MS2 was analyzed with a resolution of 30,000 at 1.6 m/z. The AGC target, MIT, and HCD collision energies were set to standard, auto, and 30%, respectively. Additionally, the sceHCD mode was enabled with an energy difference of ±10% (20-30-40%).

Data analysis and bioinformatics

Byonic software (version 3.10.10, Protein Metrics, Inc.) was used to search the RAW data files. The human IgG Uniprot database was selected. Precursor and fragment ion mass tolerances were set as ±6 ppm and ±20 ppm, respectively. The fragmentation type was set to "Both HCD & EThcD". Two missed cleavage sites were allowed. Carbamidomethyl (C) was set as a fixed modification, while Oxidation (M) and Acetyl (Protein N-term) were set as variable modifications. The N-glycan modification was set as the "182 human N-glycans". Protein groups were filtered to 1% FDR. Quality control methods for identifying intact N-glycopeptides included a Byonic score of over 200 and at least 5 amino acids. Each spectrum of intact N-glycopeptide should be confirmed manually. Abnormal data were initially excluded based on the spectral shape and TIC information. The remaining data was classified into five groups (HC, PDAC I, PDAC II, PDAC III, and PDAC IV) according to clinical information.

The intact N-glycopeptides were quantified using PANDA software (v.1.2.5) in label-free quantification mode with or without the match between runs (MBR) function. Both MS raw files and "output.spectra" files (intact glycopeptide identification files obtained from the Byonic search results) were imported. Default values were used for all other parameters. The batch search results were compiled into a file named "Peptides_F1" using PANDA. The file can be used for qualitative comparison without MBR. To compare the detection rate of N-glycopeptide between HC and PDAC patients, those glycopeptides whose detection rate ≤ 50% in both of the two groups were removed, and Chi-square was employed to calculate the p-value. For remaining glycopeptides, MBR was selected to supplement missing quantitative values. In addition, the remaining missing values were filled with the minimum value. Furtherly, the data of each column was normalized (the ratio of glycopeptide expression to total expression per column can be calculated as the glycopeptide ratio). Finally, glycopeptides were accumulated by site-specificity for further analysis.

There were two groups, HC and PDAC. The study compared the expression level of four glycopeptides subclasses and intact N- glycopeptides obtained by site-specific accumulation in the two groups using grouped boxplots. Hypothesis testing was then used to compare the expression of glycopeptides between the two groups in the Wukong platform. Statistical significance was defined as $p<0.05$. FC (fold change) >1.5 indicates up-regulated expression, while FC<0.67 indicates down-regulated expression. The boxplots with statistical significance were marked. The correlation between differential glycopeptides and clinical information including serum CA19-9 level, BMI, and age was analyzed using the Spearman correlation method in Origin (2021). The patients were divided into two groups based on gender and jaundice. The t-test ($p<0.05$) and fold change (FC > 1.5 or < 0.67) were used to determine whether there was a difference in the expression of differential glycopeptides between two groups. Binary logistic regression was used to perform univariate and multivariate analyses in SPSS (v.27) to establish a diagnostic model for PDAC patients compared to HC. Besides, K-fold cross validation was used to evaluate the model (k = 10). Subsequently, the multivariate diagnosis of pancreatic cancer was performed using receiver operating characteristic (ROC) analysis. Patients were followed up until January 31st, 2024. The Cox proportional hazards model was employed to analyze factors associated with prognosis. Statistical significance was defined as $p<0.05$.

Results

Basic characteristics of the included population

A total of 130 plasma samples were collected, consisting of 30 healthy individuals and 100 patients with PDAC at stages I (PDAC I=30), II (PDAC II=15), III (PDAC III=21), and IV (PDAC IV=34). As shown in Table 1, the average age of the two groups was 62.38 ± 10.95 years and 61.43 ± 8.89 years, respectively, while the average BMI of the two groups was 23.51 ± 2.04 and 23.08 ± 2.11, respectively. For most of the patients, the tumors ranged in size from 2 to 4 cm, and their serum CA19-9 levels were between 30 and 300 U/mL. A small number of patients were diagnosed with jaundice. The primary tumor location was most frequently in the pancreatic head. There was no statistical difference on the baseline data of the two groups, except for the CA19-9 value and jaundice (seen Additional file 1).

Qualitative analysis of intact N-glycopeptides from plasma IgGs

Based on our GlycoQuant method, a total of 94 intact N-glycopeptides from plasma IgGs were identified (seen Additional file 2). To ensure the reliability of subsequent analysis, only intact N-glycopeptides with a quantitative deletion value of less than 50% in both groups (HC and PDAC) were retained ($n=32$). The detection rates of these intact N-glycopeptides in HC and PDAC were

Table 1 Demographic characteristics of the two groups of individuals included

Variate		PDAC	HC	*p* (T-test)
Age (years)		62.38 ± 10.95	61.43 ± 8.89	0.070
BMI		23.51 ± 2.04	23.08 ± 2.11	0.994
Gender	Male	54	14	0.481
	Female	46	16	
Chronic disease	Hypertension	33	11	0.710
	Diabetes	25	6	0.573
Living habit	smoking	49	14	0.823
	Drinking	49	14	0.823
Size (cm)	≤ 2	12		
	2 < d ≤ 4	57		
	> 4	31		
CA19-9 (U/mL)	≤ 30	19	30	< 0.001
	30 < CA19-9 ≤ 300	45	0	
	300 < CA19-9 ≤ 1,000	33	0	
	> 1,000	3	0	
Jaundice	None	72	30	0.001
	Yes	28	0	
Location	head	45		
	Uncinate process	11		
	Neck	8		
	Body	25		
	Tail	11		
AJCC stage	I	30		
	II	15		
	III	21		
	IV	34		
Number of samples		100	30	

PDAC, pancreatic ductal adenocarcinoma; HC, healthy control; BMI, body mass index; CA19-9, carbohydrate antigen 19−9; AJCC, American joint committee on cancer

compared using a Chi-square test, with statistical significance defined as $p<0.05$. Finally, a total of 15 intact N-glycopeptides exhibited statistical differences between the two groups (seen Additional file 3). The file includes the peptide sequences, glycan compositions and *P* value (Chi-square test).

Quantitative analysis of intact N-glycopeptides from plasma IgGs

The updated PANDA software (v.1.2.5) was used to quantify intact N-glycopeptides of plasma IgGs in label-free quantification mode with the match between runs (MBR) function. After filling with minimum values and normalization, a total of 32 intact N-glycopeptides were quantified (seen Additional file 4). Following the removal of redundancy, 22 site-specific N-glycopeptides with accession numbers were determined for further analysis (Table 2).

The results showed no difference in expression when comparing the total expression of four subclasses of IgG in both HC and PDAC groups (Fig. 1A). In addition, to investigate the N-glycosylation variations associated with PDAC, we compared the 22 site-specific N-glycopeptides from the HC group with those from the PDAC group. Four N-glycopeptides (IgG1-GP05, IgG2-GP01, IgG3-GP03, IgG4-GP04) showed differential expression, and only IgG2-GP01 (IgG2-EEQFNSTFR-HexNAc [3]Hex [3] Fuc [1]) showed high expression in PDAC group significantly (FC = 1.71, $p<0.001$) (Fig. 1B and C).

Clinical correlation analysis

The study utilized the Spearman correlation method to examine the correlation between these four site-specific N-glycopeptides and the clinical information including age, serum CA19-9 levels, and BMI of the people enrolled (Fig. 2A). The results revealed a strong positive correlation between IgG3-GP03 and IgG4-GP04 (0.76). Additionally, there were moderate negative correlations between IgG2-GP01 and IgG1-GP05, IgG3-GP03, and IgG4-GP04, with correlation coefficients of -0.49, -0.49, and −0.46, respectively. Besides, the T-test analysis results showed no significant difference in the expression levels of differential glycopeptides between the two gender groups. However, there was a difference in the expression of IgG1-GP05 between the jaundice and non-jaundice groups. Specifically, individuals

Table 2 Quantified site-specific N-glycopeptides from IgGs

Accession number	Sequence	*N*-glycan composition
IgG1-GP01	EEQYNSTYR	HexNAc(4)Hex(3)Fuc(1)
IgG1-GP02	EEQYNSTYR	HexNAc(4)Hex(4)
IgG1-GP03	EEQYNSTYR	HexNAc(4)Hex(4)Fuc(1)
IgG1-GP04	EEQYNSTYR	HexNAc(4)Hex(5)Fuc(1)
IgG1-GP05	TKPREEQYNSTYR	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)
IgG1-GP06	EEQYNSTYR	HexNAc(5)Hex(4)Fuc(1)
IgG2-GP01	EEQFNSTFR	HexNAc(3)Hex(3)Fuc(1)
IgG2-GP02	EEQFNSTFR	HexNAc(4)Hex(3)Fuc(1)
IgG2-GP03	EEQFNSTFR	HexNAc(4)Hex(4)Fuc(1)
IgG2-GP04	EEQFNSTFR	HexNAc(4)Hex(4)Fuc(1)NeuAc(1)
IgG2-GP05	EEQFNSTFR	HexNAc(4)Hex(5)Fuc(1)
IgG2-GP06	EEQFNSTFR	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)
IgG2-GP07	EEQFNSTFR	HexNAc(5)Hex(3)Fuc(1)
IgG2-GP08	EEQFNSTFR	HexNAc(5)Hex(4)Fuc(1)
IgG3-GP01	EEQYNSTFR	HexNAc(4)Hex(3)Fuc(1)
IgG3-GP02	EEQYNSTFR	HexNAc(4)Hex(4)Fuc(1)
IgG3-GP03	EEQYNSTFR	HexNAc(4)Hex(5)Fuc(1)
IgG3-GP04	EEQYNSTFR	HexNAc(5)Hex(3)Fuc(1)
IgG4-GP01	EEQFNSTYR	HexNAc(4)Hex(3)Fuc(1)
IgG4-GP02	EEQFNSTYR	HexNAc(4)Hex(4)Fuc(1)
IgG4-GP03	TKPREEQFNSTYR	HexNAc(4)Hex(4)Fuc(1)
IgG4-GP04	EEQFNSTYR	HexNAc(4)Hex(5)Fuc(1)NeuAc(1)

without jaundice had a higher expression of IgG1-GP05 ($p<0.001$) by 2.38 times compared to those diagnosed with jaundice. (seen Additional file 5 and Additional file 6).

Regression analysis and ROC between HC and PDAC group

As demonstrated in the univariate analysis, CA19-9, IgG1-GP05, IgG2-GP01, IgG3-GP03, and IgG4-GP04 have an impact on the diagnosis. The multivariate analysis shows that the final model includes CA19-9, IgG1-GP05, and IgG4-GP04 (Table 3). The Chi-square of the Hosmer and Lemeshow Test was 0.246, with a *p*-value of 1.00>0.05, indicating that the model was well established. The ROC curve for multivariate diagnosis is presented below. The area under the curve (AUC) is 0.988 the maximum Jorden index is 0.937, the sensitivity is 0.970, the specificity is 0.967, and the optimal cut-off value is 0.616 (Fig. 2B). The results of cross-validation showed that RMSE was 0.332, R^2 was 0.428, and MAE was 0.248, indicating that the model had a good diagnostic efficiency. It is regrettable that the results of the univariate analysis of the Cox proportional hazards model, as presented in the Additional file 7, did not identify any prognostic factors related to pancreatic cancer.

Discussion

Compared to the genome and proteome which have a template, glycans are produced through a complex non-template-driven process. This process involves the competition of enzymes that extend the nascent chain, with the addition of each monosaccharide dictated by the one before. Additionally, there is no "completed" structure, as the glycans may exit anywhere along the glycosylation pathway, producing a range of structures that vary in terms of linkages, length, number of antennae, and composition [27]. The analysis of glycoproteomic is more challenging than that of other biopolymers due to the importance of selecting the appropriate fragment mode, which directly impacts the accuracy and quantity of intact glycopeptide recognition. The traditional fragmentation method sceHCD or EThcD has been used in many studies [28]. We proposed a new mass spectrometry fragmentation method, EThcD-sceHCD, which offers the benefits of both approaches and provides a higher spectral quality and identification depth compared with the traditional methods [29, 30]. Besides, we proposed an innovative approach named GlycoQuant, which integrated fast IgG preparation method, EThcD-sceHCD and updated PANDA software tool together to enable the quantification of intact N-glycopeptides [25].

IgG is the most abundant class of immunoglobulins in the circulatory system. It plays an important role in antigen neutralization, phagocytosis, cytotoxicity, and maintenance of immune homeostasis. The pattern and level of N-glycosylation in IgG Fc fragments determine most of these IgG-mediated protective immune responses. Changes in N-glycosylation are associated with different diseases, such as autoimmune diseases, infections, and cancer. Our team has demonstrated the reliability and practicability of our proposed approach, EThcD-sceHCD and GlycoQuant in a variety of research of different disease. For instance, we analyzed tryptic peptides without enrichment of intact N-glycopeptides of human plasma IgG from 58 healthy people and 111 patients with chronic kidney disease via EThcD-sceHCD. Finally, the new approach resulted in higher spectral quality, more informative fragment ions, higher Byonic score, and almost twice the depth of intact N-glycopeptide identification than sceHCD or EThcD alone. Besides, the results of the study also revealed that intact N-glycopeptides were differentially expressed in healthy people and chronic kidney disease patients, which therefore, can be a diagnostic tool [23]. In another study, we identified site-specific N-glycosylation of the human immunodeficiency virus type 1 (HIV-1) Env protein gp120 via EThcD-sceHCD. It has proven to increase the number of identified glycopeptides compared with EThcD, and produce more fragment ions than sceHCD [30]. Additionally, in a recent study, we combined EThcD-sceHCD and an intact N-glycopeptide batch quantifcation software tool (the upgraded PANDA v.1.2.5) together to analyze intact N-glycopeptides of plasma IgGs from 58 healthy people, 48 patients with membranous nephropathy, and 35 patients with IgA

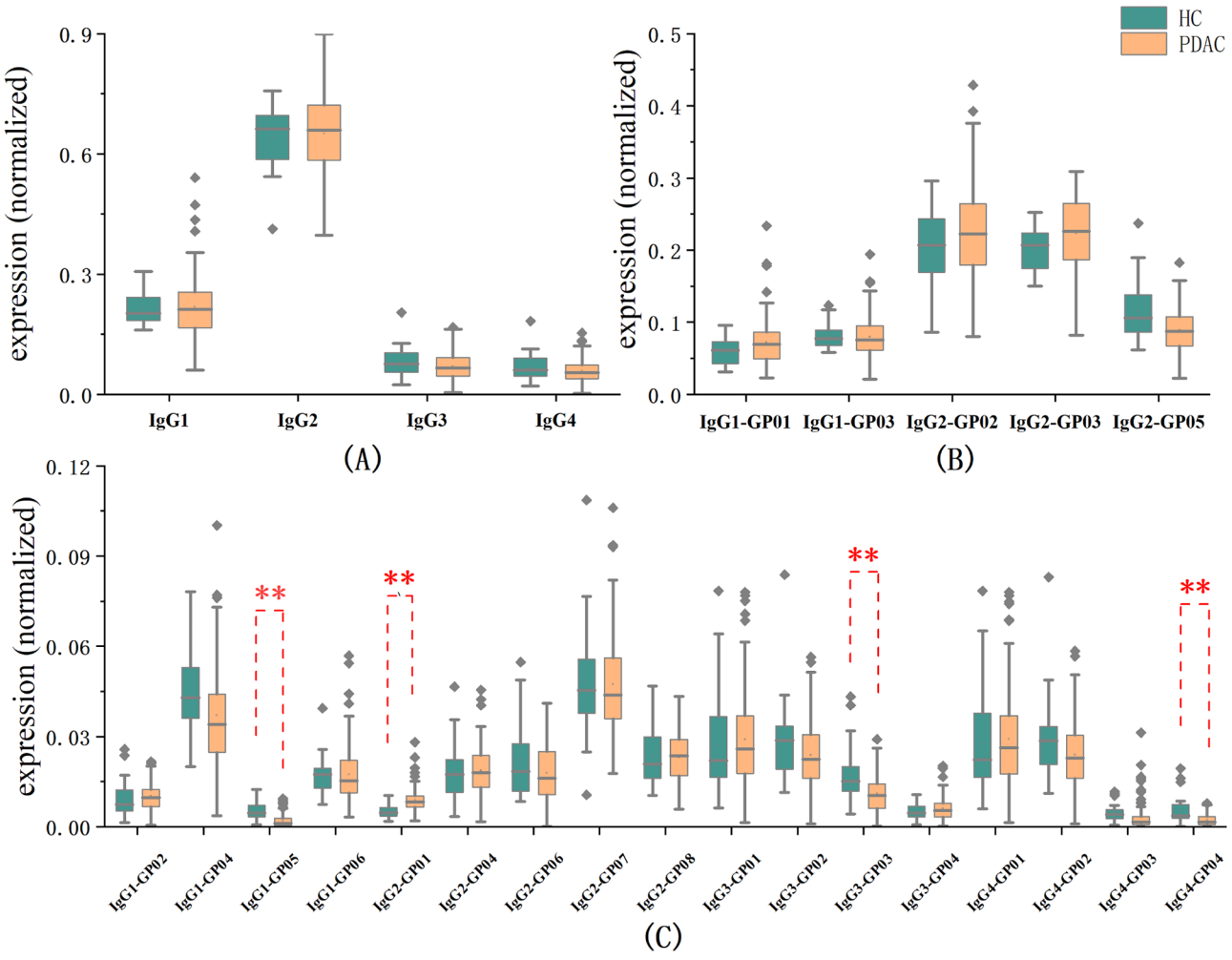

Fig. 1 Differentially expressed site-specific N-glycopeptides between HC and PDAC groups. (**A**), differentially expressed N-glycopeptides subclasses between HC and PDAC groups; (**B**) and (**C**), differentially expressed site-specific N-glycopeptides between HC and PDAC groups; HC, healthy control; PDAC, pancreatic ductal adenocarcinoma; IgG, immunoglobulin G; *, $p<0.05$; **, $p<0.01$

nephropathy. The final results identified several quantified intact N-glycopeptides which could distinguish not only between chronic kidney disease and healthy people, but also between different types of chronic kidney disease (membranous nephropathy and IgA nephropathy) and served as a diagnostic tool for renal tubular function [25]. These researches demonstrated the feasibility of the new approach we proposed and as well, reveal the potential of intact N-glycopeptides of IgG for disease diagnosis.

IgG glycosylation has also been investigated as a potential source of markers for pancreatic cancer. A study analyzed the serum IgG-glycosylation from 86 autoimmune pancreatitis patients, 115 PDAC patients and 57 healthy people. The study suggested that IgGs had different rations of glycosylation type between autoimmune pancreatitis patients and PDAC patients, with a higher agalactosylation ratio of IgG1 and a higher agalactosylation ratio of IgG2 for PDAC patients, and a higher sialylation ratio of IgG subclasses 1, 2 and 4 for autoimmune pancreatitis patients. However, this study did not explore specific glycosylation patterns, nor did it involve quantitative studies of specific glycopeptide expression [31]. Additionally, another study analyzed six fucosylated glycoforms of IgG2 in the plasma of 75 healthy people, 75 pancreatitis patients and 75 PDAC patients. The results showed significant increases in agalactosylated glycoforms and decreases in galactosylated glycoforms for pancreatitis and PDAC patients compared with healthy controls [32]. However, the study just focused on the IgG2, and similarly, didn't involve quantitative studies of specific glycopeptide expression. In addition, another study utilized a novel method to determine the glycosylation of IgG subclasses. They identified seven glycopeptides that have high potential in differentiating between autoimmune pancreatitis and PDAC. The glycopeptides that identified in the study were not the same as ours, which maybe because that the study didn't point out the specific glycosylation pattern, N-glycosylation or

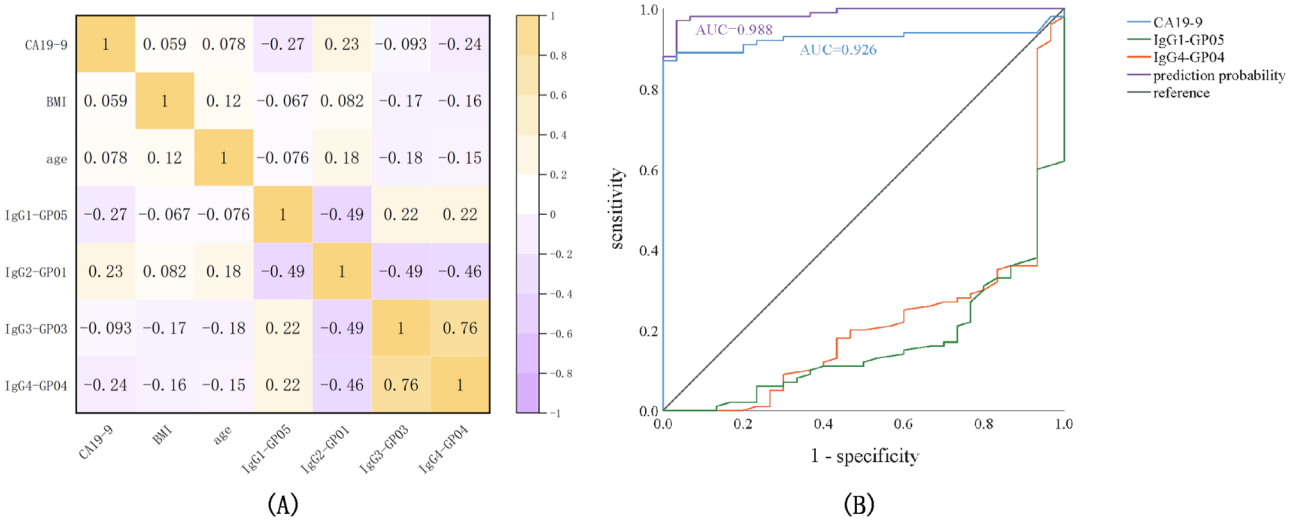

Fig. 2 Clinical correlation analysis including Spearman correlation analysis and ROC of PDAC diagnosis. (**A**), Spearman correlation analysis of serum CA19-9 levels, BMI, age and four site-specific N-glycopeptides, the yellow color represents a positive correlation, while the purple color represents a negative correlation. The different depth of the color represents the different value of the number, which ranges from −1 and 1. The deeper the color, the larger the absolute value; (**B**), ROC of CA19-9, IgG1-GP05, and IgG4-GP04 combined diagnosis for patients with PDAC; binary logistics regression was used for ROC analysis. HC, healthy control; PDAC, pancreatic ductal adenocarcinoma; IgG, immunoglobulin G; CA199, carbohydrate antigen 19−9; BMI, body mass index; ROC, receiver operating characteristic

Table 3 Univariate and multivariate analysis of diagnosis between HC and PDAC group

Variate	Univariate analysis		Multivariate analysis		
	HR (95% CI)	*p*	β	HR (95% CI)	*p*
age	1.01 (0.97–1.05)	0.664	-	-	-
Gender	1.34 (0.59–3.04)	0.481	-	-	-
BMI	1.11 (0.91–1.36)	0.307	-	-	-
Hypertension	0.85 (0.36–1.99)	0.710	-	-	-
Diabetes	1.33 (0.49–3.63)	0.574	-	-	-
Smoking	1.10 (0.49–2.49)	0.823	-	-	-
Drinking	1.10 (0.49–2.49)	0.823	-	-	-
CA19-9	1.14 (1.06–1.23)	<0.001	0.12	1.12 (1.04–1.22)	0.005
IgG1-GP05	0.02 (0.004–0.09)	<0.001	-7.61	0.00 (0.00-0.15)	0.009
IgG2-GP01	44.72 (7.42–269.40)	<0.001	-	-	-
IgG3-GP03	0.30 (0.16–0.58)	<0.001	-	-	-
IgG4-GP04	0.012 (0.001–0.097)	<0.001	-6.24	0.00(0.00-0.63)	0.034

HR, hazard ratio; BMI, body mass index; CA19-9, carbohydrate antigen 19−9

O-glycosylation or others. Besides, the study involved in 30 clinical samples, 10 PDAC patients, 10 autoimmune pancreatitis patients and 10 healthy people, which may need more samples [33]. Compared with these researches mentioned above, our research conducted a more detailed and comprehensive investigation into IgG intact N-glycopeptides and their correlation with clinical information for the first time.

We identified four intact N-glycopeptides that showed statistically difference between PDAC patients and healthy individuals. Among the four N-glycopeptides, IgG2-GP01 has a different expression trend in patients with PDAC than the other three, which is consistent with the result of Spearman correlation analysis that IgG2-GP01 had a negative correlation with the other three N-glycopeptides. Furthermore, the AUC values for the multifactor combined diagnosis ROC curve were 0.988 for patients with PDAC, which is higher than CA19-91 along, 0.926. Besides, the result of cross validation showed that the diagnosis model had a good efficiency. This satisfactory result suggests that IgG1-GP05 and IgG4-GP04 have the potential to assist CA19-9 in PDAC diagnosis. What's more, clinical correlation analysis didn't show that IgG1-GP05 or IgG4-GP04 had correlation with clinical information except for jaundice, which indicated that these glycopeptides were stable indicators that were not influenced by clinical factors. We didn't identify any prognostic factors related to

pancreatic cancer. It may be because that the study of pancreatic cancer prognosis is a multifaceted and complex process. It is influenced not only by the stage and severity of the disease itself but also by various external factors including the surgeon's surgical technique, the surgical approach employed, post-operative chemotherapy, genetic mutations of the patient, and even the patient's psychological state. Given the numerous variables involved, future research on prognosis should aim to control for these confounding factors to ensure more reliable and precise outcomes.

There are also some limitations in our research. Larger sample size is needed for verification of the four differentially expressed intact N-glycopeptides we identified. Additionally, our study only examined the expression of IgG intact N-glycopeptide in HC and patients with PDAC. Future studies should include more samples from other pancreatic diseases, such as pancreatitis, benign and borderline pancreas tumors.

Conclusion

Our study used a well-evaluated approach (termed GlycoQuant) for the site-specific N-glycosylation profile of human plasma IgG in both HC and PDAC groups. There were four intact N-glycopeptides were found differentially expressed in HC and patients with PDAC. These findings provide potential target N-glycopeptides for further studies with larger sample size. The ROC curve obtained in our study offers a new approach for the diagnosis of PDAC. However, our study was unable to identify the exact factors that affect the prognosis of PDAC due to the diverse range of factors, such as surgical techniques, postoperative complications, and patients' physical conditions. Further research is required to investigate this.

Abbreviations

PDAC	Pancreatic ductal adenocarcinoma
CA19-9	Carbohydrate antigen199
IgG	Immunoglobulins G
EthcD	Electron transfer high energy collision dissociation
SceHCD	Stepped collision energy high energy collision dissociation
EThcD-SceHCD	Electron transfer high energy collision dissociation-Stepped collision energy high
HC	Healthy control
BMI	Body mass index
TNM	Tumor–node–metastasis
AJCC	American joint committee on cancer
ABC	Ammonium bicarbonate
MBR	Match between runs
FC	Fold change
ROC	Receiver operating characteristic
AUC	Area under the curve

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12014-024-09522-4.

Supplementary Material 1: Clinical information and N-glycopeptide data for all people enrolled

Supplementary Material 2: A total of 94 glycopeptides that were initially identified

Supplementary Material 3: A total of 15 glycopeptides that have differential detection rates in PDAC and HC

Supplementary Material 4: Quantitative data of 32 glycopeptides that were identified after missing process

Supplementary Material 5: Clinical correlation analysis between the four glycopeptides and gender

Supplementary Material 6: Clinical correlation analysis between the four glycopeptides and jaundice

Supplementary Material 7: Univariate analysis of prognosis for patients with PDAC

Acknowledgements

Not Applicable.

Author contributions

Yong Zhang and Ang Li designed the research; Yi Jin, Ran Hu, Yufan Gu and Ailin Wei, processed the samples and performed LC-MS/MS analysis; Yong Zhang and Yi Jin performed data analysis and wrote the manuscript. All authors have approved the final version of the manuscript.

Funding

This work was supported by the National Key R&D Program of China (2022YFF0608401), and the Department of Science and Technology of Sichuan Province (2021YFS0109).

Data availability

The RAW data and search results were deposited in the ProteomeXchange Consortium via the iProX partner repository, with the dataset identifier PXD051436.

Declarations

Ethics approval and consent to participate

The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of West China Hospital of Sichuan University (No. 2022 – 597). All experiments meet the ethical requirements, and all data are true and reliable.

Consent for publication

Not Applicable.

Competing interests

The authors declare no competing interests.

Received: 23 September 2024 / Accepted: 16 December 2024
Published online: 30 December 2024

References

1. Mizrahi JD, Surana R, Valle JW, Shroff RT. Pancreatic cancer. Lancet. 2020;395(10242):2008–20.
2. Siegel RL, Miller KD, Fuchs HE, Jemal A, Cancer, Statistics. 2021. Ca-a Cancer Journal for Clinicians. 2021;71(1):7–33.
3. Sarantis P, Koustas E, Papadimitropoulou A, Papavassiliou AG, Karamouzis MV. Pancreatic ductal adenocarcinoma: treatment hurdles, tumor microenvironment and immunotherapy. World J Gastrointest Oncol. 2020;12(2).
4. Hartwig W, Hackert T, Hinz U, Gluth A, Bergmann F, Strobel O, et al. Pancreatic Cancer surgery in the New Millennium Better Prediction of Outcome. Ann Surg. 2011;254(2):311–9.

5. Huang J, Lok V, Ngai CH, Zhang L, Yuan J, Lao XQ, et al. Worldwide Burden of, risk factors for, and trends in Pancreatic Cancer. Gastroenterology. 2021;160(3):744–54.
6. Wong CH. Protein glycosylation: new challenges and opportunities. J Org Chem. 2005;70(11):4219–25.
7. Kuzmanov U, Kosanam H, Diamandis EP. The sweet and sour of serological glycoprotein tumor biomarker quantification. BMC Med. 2013;11.
8. Kailemia MJ, Park D, Lebrilla CB. Glycans and glycoproteins as specific biomarkers for cancer. Anal Bioanal Chem. 2017;409(2):395–410.
9. Munkley J, Elliott DJ. Hallmarks of glycosylation in cancer. Oncotarget. 2016;7(23):35478–89.
10. Hart GW. Myriad roles of glycans in Biology. J Mol Biol. 2016;428(16):3147–9.
11. Llop E, Guerrero PE, Duran A, Barrabes S, Massaguer A, Jose Ferri M, et al. Glycoprotein biomarkers for the detection of pancreatic ductal adenocarcinoma. World J Gastroenterol. 2018;24(24):2537–54.
12. Silva MLS. Cancer serum biomarkers based on aberrant post-translational modifications of glycoproteins: clinical value and discovery strategies. Biochim Et Biophys Acta-Reviews Cancer. 2015;1856(2):165–77.
13. Peracaula R, Barrabes S, Sarrats A, Rudd PM, de Llorens R. Altered glycosylation in tumours focused to cancer diagnosis. Dis Markers. 2008;25(4–5):207–18.
14. Luo G, Jin K, Deng S, Cheng H, Fan Z, Gong Y et al. Roles of CA19-9 in pancreatic cancer: Biomarker, predictor and promoter. Biochim Et Biophys Acta-Reviews Cancer. 2021;1875(2).
15. Hess V, Glimelius B, Grawe P, Dietrich D, Bodoky G, Ruhstaller T, et al. CA 19 – 9 tumour-marker response to chemotherapy in patients with advanced pancreatic cancer enrolled in a randomised controlled trial. Lancet Oncol. 2008;9(2):132–8.
16. Zhang D, Hou W, Liu F, Yin J, Lu W, Li M, et al. Metformin reduces serum CA199 levels in type 2 diabetes Chinese patients with time-effect and gender difference. Diabetes Technol Ther. 2015;17(2):72–9.
17. Kim S, Park BK, Seo JH, Choi J, Choi JW, Lee CK et al. Carbohydrate antigen 19 – 9 elevation without evidence of malignant or pancreatobiliary diseases. Sci Rep. 2020;10(1).
18. Indellicato R, Zulueta A, Caretti A, Trinchera M. Complementary use of Carbohydrate Antigens Lewis a, Lewis b, and Sialyl-Lewis a (CA19.9 Epitope) in gastrointestinal cancers: Biological Rationale towards a personalized clinical application. Cancers (Basel). 2020;12(6).
19. Sato Y, Fujimoto D, Uehara K, Shimizu R, Ito J, Kogo M et al. The prognostic value of serum CA 19 – 9 for patients with advanced lung adenocarcinoma. BMC Cancer. 2016;16.
20. Ansari D, Toren W, Zhou Q, Hu D, Andersson R. Proteomic and genomic profiling of pancreatic cancer. Cell Biol Toxicol. 2019;35(4):333–43.
21. Ritamo I, Cloutier M, Valmu L, Neron S, Rabina J. Comparison of the glycosylation of in vitro generated polyclonal human IgG and therapeutic immunoglobulins. Mol Immunol. 2014;57(2):255–62.
22. Zhang Y, Lin T, Zhao Y, Mao Y, Tao Y, Huang Y, et al. Characterization of N-linked intact glycopeptide signatures of plasma IgGs from patients with prostate carcinoma and benign prostatic hyperplasia for diagnosis pre-stratification. Analyst. 2020;145(15):5353–62.
23. Zhang Y, Zheng S, Mao Y, Cao W, Zhao L, Wu C, et al. Systems analysis of plasma IgG intact N-glycopeptides from patients with chronic kidney diseases via EThcD-sceHCD-MS/MS. Analyst. 2021;146(23):7274–83.
24. Liu J, Dolikun M, Stambuk J, Trbojevic-Akmacic I, Zhang J, Zhang J, et al. Glycomics for type 2 diabetes Biomarker Discovery: Promise of Immunoglobulin G Subclass-Specific Fragment Crystallizable N-glycosylation in the Uyghur Population. Omics-a J Integr Biology. 2019;23(12):640–8.
25. Zhao Y, Zhang Y, Meng B, Luo M, Li G, Liu F et al. A Novel Integrated Pipeline for Site-Specific quantification of N-glycosylation. Phenomics. 2024.
26. Yang YM. Evaluation of American Joint Commission on Cancer(8th edition) and Japanese Pancreas Society(7th edition)changes for T and N staging in patients with pancreatic adenocarcinoma. Zhonghua Wai Ke Za Zhi [Chinese J Surgery]. 2017;55(1):20–3.
27. Ruhaak LR, Xu G, Li Q, Goonatilleke E, Lebrilla CB. Mass Spectrometry approaches to Glycomic and glycoproteomic analyses. Chem Rev. 2018;118(17):7886–930.
28. Riley NM, Malaker SA, Driessen MD, Bertozzi CR. Optimal dissociation methods differ for N- and O-Glycopeptides. J Proteome Res. 2020;19(8):3286–301.
29. Luo M, Mao Y, Zeng W, Zheng S, Li H, Hu J et al. Site-specific N-glycosylation characterization of micro monoclonal immunoglobulins based on EThcD-sceHCD-MS/MS. Front Immunol. 2022;13.
30. Zhang Y, Zheng S, Zhao W, Mao Y, Cao W, Zeng W et al. Sequential analysis of the N/O-Glycosylation of heavily glycosylated HIV-1 gp120 using EThcD-sceHCD-MS/MS. Front Immunol. 2021;12.
31. Shih H-C, Chang M-C, Chen C-H, Tsai IL, Wang S-Y, Kuo Y-P et al. High accuracy differentiating autoimmune pancreatitis from pancreatic ductal adenocarcinoma by immunoglobulin G glycosylation. Clin Proteom. 2019;16.
32. Chen G, Li H, Qiu L, Qin X, Liu H, Li Z. Change of fucosylated IgG2 Fc-glycoforms in pancreatitis and pancreatic adenocarcinoma: a promising disease-classification model. Anal Bioanal Chem. 2014;406(1):267–73.
33. Shiao J-Y, Chang Y-T, Chang M-C, Chen MX, Liu L-W, Wang X-Y et al. Development of efficient on-bead protein elution process coupled to ultra-high performance liquid chromatography-tandem mass spectrometry to determine immunoglobulin G subclass and glycosylation for discovery of bio-signatures in pancreatic disease. J Chromatogr A. 2020;1621.

Publisher's note