



Universiteit  
Leiden  
The Netherlands

## Glyco(proteo)mic workflows for cancer biomarker discovery

Moran, A.B.

### Citation

Moran, A. B. (2023, November 1). *Glyco(proteo)mic workflows for cancer biomarker discovery*. Retrieved from <https://hdl.handle.net/1887/3655862>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/3655862>

**Note:** To cite this publication please use the final published version (if applicable).

# Chapter 5

## Serum N-Glycosylation RPLC-FD-MS Assay to Assess Colorectal Cancer Surgical Interventions

Alan B. Moran<sup>1,2</sup>, Georgia Elgood-Hunt<sup>2</sup>, Yuri E.M. van der Burgt<sup>1</sup>, Manfred Wuhrer<sup>1</sup>, Wilma E. Mesker<sup>3</sup>, Rob A.E.M. Tollenaar<sup>3</sup>, Daniel I.R. Spencer<sup>2</sup> and Guinevere S.M. Lageveen-Kammeijer<sup>1,4</sup> \*

<sup>1</sup> Leiden University Medical Center, Center for Proteomics and Metabolomics, 2300 RC Leiden, The Netherlands

<sup>2</sup> Ludger Ltd., Culham Science Centre, OX14 3EB Abingdon, United Kingdom

<sup>3</sup> Leiden University Medical Center, Department of Surgery, 2300 RC Leiden, The Netherlands

<sup>4</sup> University of Groningen, Department of Analytical Biochemistry, Groningen Research Institute of Pharmacy, Groningen, The Netherlands

Reprinted (adapted) from Moran, A. B., Elgood-Hunt, E., van der Burgt Y. E. M., Wuhrer, M., Mesker W. E., Tollenaar, R. A. E. M., Spencer, D. I. R., Lageveen-Kammeijer, G. S. M. Serum N-glycosylation RPLC-FD-MS Assay to Assess Colorectal Cancer Surgical Interventions. *Biomolecules*. **2023**, 13 (6), 896, <https://doi.org/10.3390/biom13060896> under the Creative Commons CC-BY license.



## Abstract

A newly developed analytical strategy was applied to profile the total serum *N*-glycome of 64 colorectal cancer (CRC) patients before and after surgical intervention. In this cohort it was previously found that serum *N*-glycome alterations in CRC associated with patient survival. Here, fluorescent labeling of serum *N*-glycans was applied using procainamide and followed by sialic acid derivatization specific for  $\alpha$ 2,6- and  $\alpha$ 2,3-linkage types via ethyl esterification and amidation, respectively. This strategy allowed efficient separation of specific positional isomers on reversed-phase liquid chromatography–fluorescence detection–mass spectrometry (RPLC-FD-MS) and complemented the previous glycomics data based on matrix-assisted laser desorption/ionization (MALDI)-MS that did not include such separations. The results from comparing pre-operative CRC to post-operative samples were in agreement with studies that identified a decrease of di-antennary structures with core fucosylation and an increase in sialylated tri- and tetra-antennary *N*-glycans in CRC patient sera. Pre-operative abundances of *N*-glycans showed good performance for the classification of adenocarcinoma and led to the revisit of the previous MALDI-MS dataset with regard to histological and clinical data. This strategy has potential to monitor patient profiles before, during and after clinical events such as treatment, therapy or surgery and should also be further explored.

## Introduction

Glycans drastically affect the function of a glycoprotein and contribute to the malignant phenotype of cancer cells by promoting proliferation, metastasis, and immunosuppression.<sup>1–3</sup> In this regard, aberrant glycosylation profiles add to understanding pathological steps of tumor development and progression and have become a hallmark of cancer.<sup>4,5</sup> For example, in the case of colorectal cancer (CRC), aberrant *N*-glycosylation with regard to tumor microenvironment is increasingly studied in CRC tissue samples using imaging mass spectrometry (MS).<sup>6–9</sup> Other glycomics studies also reported associations of *N*-glycome alterations with survival and tumor stage in CRC.<sup>10–12</sup> Overall, there is evidence to support that earlier diagnosis of cancer increases success of curative treatment and long-term survival in general and has motivated various omics-fields to explore retrospectively collected serum samples from cancer patients.<sup>13</sup>

Previously we have applied matrix-assisted laser desorption/ionization (MALDI)-MS as a readout for total serum *N*-glycome analysis from CRC patients where it was found that serum *N*-glycome alterations in CRC associated with patient survival.<sup>14</sup> Although MALDI-MS is a fast detection technique, it does not include any separation of positional isomers.<sup>15</sup> In contrast, recently we developed a platform whereby plasma *N*-glycans are fluorescently labeled using procainamide followed by sialic acid linkage-specific derivatization via ethyl esterification and amidation which allows efficient separation of sialylated and positional isomers on reversed-phase liquid chromatography–fluorescence detection–mass spectrometry (RPLC-FD-MS).<sup>16</sup> The performance of this assay was reported for plasma samples, but this method is also suitable for serum-based specimen. In particular, positional isomers which are related to galactose occupation on the glycan arm ( $\alpha 3$  versus  $\alpha 6$ ) as well as fucose location (antennary versus core) were separated. Thus, we reported this effective analytical strategy by benchmarking the performance of a complex mixture of released plasma *N*-glycans against the gold standard; hydrophilic interaction liquid chromatography (HILIC)-FD-MS.<sup>16</sup> We demonstrated sufficient complementarity between both platforms as an overlap up to 95% of the total relative area was demonstrated with the HILIC-FD-MS approach. Moreover, the repeatability and intermediate precision of the RPLC-FD-MS platform were determined at a median interday ( $n = 50$ ) relative standard deviation (RSD) of 5.4% for the 10 most abundant *N*-glycans. The overall

throughput, compared to a HILIC-approach for released *N*-glycan analysis, was improved as the sample preparation protocol is executed on a liquid handling robot and a 35 minute separation gradient is applied.

In the current work we re-visit the previous MALDI-MS-based serum glycomics study in order to demonstrate the clinical validity of the developed method,<sup>14</sup> and apply the RPLC-FD-MS platform on a sub-set of the samples, namely pre- and post-operative serum samples obtained from 64 CRC patients. In particular, we aim to uncover relevant serum *N*-glycomic signatures in the context of CRC, specifically related to sialic acid linkages as well as positional isomeric structures. The results will be compared with those from the MALDI-MS-based measurements of the pre- and post-operative samples and furthermore discussed in the context of histological type and cases versus controls differences.<sup>14</sup>

## Materials and Methods

### Study Design

Pre- and post-operative serum samples were collected from 64 CRC patients during the period of October 2002 and March 2013 using a standardized protocol by the LUMC Surgical Oncology Biobank, resulting in 128 samples in total. Post-operative samples were defined as “cancer-free” based on a post-operative wash out period (45 days), absence of recurrence, and a follow-up assessment after at least 1393 days. The cohort characteristics, following data curation steps that are described in the following sections, are shown in **Table 1**. The procedure for sample collection and storage of pre- and post-operative patient serum samples was previously reported.<sup>14</sup> Serum tubes (BD Vacutainer SST II plastic serum tubes, BD, United Kingdom) were used to draw blood before being centrifuged within a maximum of four hours. Both pre-operative and post-operative samples were stored for 30 days or less at - 20 °C before biobank collection and were then stored at - 80 °C. Aliquoting was carried out using a Microlab STAR line pipetting robot (Hamilton, Switzerland) into sterile 2D barcoded V-bottom 500 µL tubes (Thermo Scientific, Hudson), prior to this samples were thawed on ice and the aliquots were stored at -80 °C. Prior to sample preparation, the samples were stored at -20 °C before being thawed in order to begin sample processing.

**Table 1. Clinical characteristics of the pre- and post-operative CRC cohort.** Samples were taken before and after surgery from the same patient. A total of 64 patients were included in the analysis following data curation. Two measurements per patient were carried out in order to process the pre- and post-operative samples. Full metadata was collected for each patient in three categories of the disease: stage (1 – 4), local extent of tumor assessed pathologically (pT1 – 4), and histological type (Signet-ring cell carcinoma, >50% signet-ring cells; adenocarcinoma, and mucinous adenocarcinoma, >50% mucinous). Standard deviation (s.d.).

Pre- and post-operative CRC cohort (n = 64)		
Female sex, n (%)		33 (51.6)
Age in years, mean (s.d.)		64.2 (12.7)
Stage, n (%)	1	13 (20.3)
	2	31 (48.4)
	3	16 (25.0)
	4	4 (6.3)
pT, n (%)	pT1	7 (10.9)
	pT2	8 (12.5)
	pT3	42 (65.6)
	pT4	7 (10.9)
Histological type, n (%)	Signet-ring cell carcinoma	3 (4.7)
	Adenocarcinoma	54 (84.4)
	Mucinous adenocarcinoma	7 (10.9)

The entire workflow, including PNGase-F release, procainamide fluorescent labeling, ethyl esterification and amidation derivatization, and RPLC-FD-MS measurement, has previously been described<sup>16</sup> and is briefly explained in the following sections. The sample preparation protocol was performed using a Hamilton Microlab STARlet liquid-handling robot.

### N-glycan Release and Fluorescent Labeling

The entire cohort of pre- and post-operative serum samples were contained in two plates. Serum (5 µL) from each sample was manually and randomly added to a 96-well PCR plate (4ti-0960/C, 4titude Ltd., United Kingdom). Importantly, the samples were distributed over two plates in total and the paired pre- and post-operative samples from the same patient were contained on the same 96-well PCR plate. Commercial plasma (P9523, Sigma-Aldrich, United Kingdom) with a concentration of 1 mg/mL (n = 26, 13 per plate) was included as a positive control, deionized water (n

= 6, 18.2 M $\Omega$  resistivity, Sartorius Arium Comfort, Germany) was processed as a negative control on plate two. This was followed by the addition of 4  $\mu$ L deionized water and 1  $\mu$ L of a 10X denaturing solution obtained from an *N*-glycan release kit [LZ-rPNGASEF-96, Ludger Ltd., United Kingdom). Following this, incubation was performed at 100 °C for 10 min. The plate was removed from the oven and allowed to cool to room temperature before deionized water (6  $\mu$ L), reaction buffer (2  $\mu$ L, 500 mM sodium phosphate, pH 7.5), and 10% NP-40 (2  $\mu$ L) were added to each well. The PNGase F solution was diluted 1:1 (*v/v*) with PNGase F storage buffer, which consisted of 50 mM sodium chloride, 5 mM ethylenediaminetetraacetic acid, and 20 mM tris-hydrochloric acid (150  $\mu$ L, New England Biolabs, United Kingdom). Then, 2  $\mu$ L was added to each sample and the plate was incubated (37 °C) overnight. The next day, a 5% formic acid (Sigma-Aldrich) / deionized water solution (*v/v*; 5  $\mu$ L) was added to each well and incubation was performed at room temperature for 40 min.

A clean-up procedure was performed using a Protein Binding Membrane (PBM) plate [LC-PBM-96, Ludger Ltd.]. The plate was prepared with successive washes of methanol (100  $\mu$ L) and deionized water (300  $\mu$ L) followed by the addition of sample to the PBM plate. The PCR plate was washed with deionized water (90  $\mu$ L) which was also added to the PBM plate. Following this, a 2 mL collection plate [LP-COLLPLATE-2ML-96, Ludger Ltd.] was placed underneath the PBM plate and elution was carried out via centrifugation at 800 rcf (22°C) for 3 min. An additional 100  $\mu$ L deionized water was added to each well and another centrifugation step was performed. The samples were transferred to a 300  $\mu$ L 96-well PCR plate [4ti-0710/C, 4titude Ltd.] and dried down using a vacuum centrifuge.

Procainamide labeling solution [LT-KPROC-96, Ludger Ltd.] was prepared according to the vendor's guidelines. The labeling solution (20  $\mu$ L) was added to each well of dried released *N*-glycans and incubated (65 °C) for 60 min. After 10 min, the plate was removed, vortexed and centrifugated before being placed back in the oven for the remaining incubation time. A HILIC clean-up plate [LC-PROC-96, Ludger Ltd.] was prepared by washing with solutions of 70% ethanol (200  $\mu$ L, Sigma-Aldrich) / deionized water (*v/v*), deionized water (200  $\mu$ L) and acetonitrile (CH<sub>3</sub>CN, 200  $\mu$ L, Romil Ltd., United Kingdom). The sample was added to the plate followed by the addition of CH<sub>3</sub>CN (230  $\mu$ L) to each well. A vacuum was applied before three further CH<sub>3</sub>CN washes (200  $\mu$ L) were carried out. The samples were eluted in 100  $\mu$ L of deionized

water into a 96-well 2 mL collection plate [LP-COLLPLATE-2 ML-96, Ludger Ltd.]. Following this, a second elution was carried out using deionized water which resulted in a final volume of 200  $\mu$ L.

### Sialic Acid Derivatization

The derivatization technique has been reported in the literature<sup>17,18</sup> and the protocol was followed as previously described.<sup>16</sup> The released and labeled *N*-glycans were transferred (120  $\mu$ L) to a 300  $\mu$ L 96-well PCR plate and dried down using a vacuum centrifuge. Following this, the samples were reconstituted in deionized water (15  $\mu$ L) and 3  $\mu$ L was added to 60  $\mu$ L of ethyl esterification solution, which consisted of 250 mM 1-ethyl-3-(3-(dimethylamino) propyl) carbodiimide (EDC, Fluorochem, United Kingdom) and 250 mM hydrate 1-hydroxybenzotriazole (HOBT, Sigma-Aldrich) dissolved in ethanol. Incubation (37 °C) was performed for 60 min before 28% ammonia (12  $\mu$ L, Sigma-Aldrich) was added to the samples and the plate was incubated for a further 60 min at 37 °C. Following this, the volume in each well was brought up to 300  $\mu$ L by the addition of 225  $\mu$ L of CH<sub>3</sub>CN. A clean-up was performed using the HILIC clean-up plate procedure described above and may also be found in the literature<sup>16</sup>. The fluorescently labeled and derivatized *N*-glycans were eluted in 200  $\mu$ L deionized water and stored at -20°C until further analysis was performed.

### Reversed-Phase Liquid Chromatography–Fluorescence Detection–Mass Spectrometry

A single workup of each sample was performed followed by single injection on the RPLC-MS-platform. The measurements for each plate were carried out separately and additional deionized water blanks ( $n = 6$ ) were included during the measurement of plate one. Samples (95  $\mu$ L) were added up to 100  $\mu$ L with CH<sub>3</sub>CN so that the final concentration was 5% CH<sub>3</sub>CN (*v/v*). A 150 x 2.1 mm, ACE excel 2 C18-PFP column (ACE Ltd., United Kingdom) was prepared on an Ultimate 3000 UHPLC system (Thermo Scientific, United Kingdom) and a 20  $\mu$ L injection was performed. The column temperature was set to 60 °C, and the fluorescence detector ( $\lambda_{\text{ex}} = 310$  nm  $\lambda_{\text{em}} = 370$  nm) bulb power was set to 'high' whilst the sensitivity was set to eight. A separation gradient was carried out using a flow rate of 0.4 mL/min and 50 mM ammonium formate as solvent A and 10% CH<sub>3</sub>CN, 0.1% formic acid (*v/v*) as solvent B. The gradient consisted of the following steps: 0 – 26.5 min (15% – 95% solvent B), 26.5 –



30.5 min (95% solvent B), 30.5 – 32.5 min (95% – 15% solvent B), 32.5 to 35.1 min (15% solvent B). A representative chromatogram with numbered peaks is shown in **Supplementary Figure S1**.

Electrospray ionization–MS was carried out on an amaZon Speed ETD MS (Bruker Daltonics GmbH, Germany). The instrument parameters were as follows: ionization mode, positive mode with enhanced resolution scanning; mass range,  $m/z$  600 – 1600 and  $m/z$  900 set as the target mass; capillary voltage, 4.5 kV; source temperature, 250 °C; gas flow, 10 L/min; max accumulation time, 50.00 ms; ion charge control (ICC) target, 200,000. MS/MS spectra were obtained using helium collision-induced dissociation under the following conditions: max accumulation time MS(n): 40.00 ms; ICC target MS(n): 200,000; precursor ion selection: 3 precursors which are released after 0.2 min; MS(n) isolation width: 4  $m/z$ ; scan range selection MS(n), scale to precursor; automatic MS(n) absolute signal threshold: 25,000; automatic MS(n) relative signal threshold: 5%.

### Data Processing

The assignment of serum *N*-glycans was performed based on our previous study.<sup>16</sup> Briefly, this was carried out based on exact mass and retention order as well as diagnostic ions obtained in the MS/MS spectra. As a result, *N*-glycans with differently linked sialic acids are mass differentiated following sialic acid derivatization.<sup>16</sup> In addition, isomers with the same mass but different retention time (RT) are denoted as “isomer 1” (iso1) and “isomer 2” (iso2) based on their relative retention order. Notably, several *N*-glycans were not detected in our previous work yet were identified in this study. In this case, the diagnostic fragment ions which support the assignment of these structures are included in **Supplementary Table S1**. *N*-glycan compositions are displayed according to the notation recommended by the Consortium for Functional Glycomics:<sup>19</sup> *N*-acetylglucosamine (N; blue square), mannose (H; green circle), galactose (H; yellow circle), fucose (F; red triangle), *N*-acetylneuraminic acid (S; purple diamond).

Fluorescence detection (FD) data was exported (.txt) using Chromeleon (version 7.2) and imported into HappyTools (version 0.1-beta1, build 190115a)<sup>20</sup> for processing. The processing parameters for this software have previously been reported<sup>16</sup>. Briefly, these parameters were as follows: datapoints (100), baseline function order (1),

background window (1), peak edge type (sigma), peak edge value (2), minimum number of calibrants (4), minimum S/N (27). Importantly, baseline correction was applied and parameters for peak calibration and integration are included in **Supplementary Table S2**. Furthermore, a peak integration window (70% FWHM) was implemented which allowed sufficient peak area quantification while also minimizing peak overlap. In addition, previously optimized quality control parameters<sup>16</sup> were applied in order to perform data curation of fluorescent chromatogram peaks. For example, the maximum RT deviation was  $\pm 3$  sec and a S/N > 9 was required. With regard to the latter, a more lenient S/N criteria in comparison with analytical validation guidelines (S/N > 10)<sup>21</sup> was implemented. Furthermore, peaks eluting at the expected RT were required to be greater than the average intensity plus nine times the standard deviation (s.d.) of the same integration window in the blank and negative control measurements. A single measurement from the analysis did not meet the quality control requirements and, therefore, both paired measurements were removed from the study, resulting in a final sample size of  $n = 64$  (128 measurements).

Raw RPLC-MS data acquired in the Bruker data format (.d) were converted into the .mzXML format and imported in LaCyTools (version 2.0, build 200723).<sup>22</sup> Glycan compositions used for spectra alignment and extraction are included in **Supplementary Table S3**. Further processing parameters used for LaCyTools have previously been described.<sup>16</sup> Importantly, the area for each charge state of each glycan composition was integrated with at least 95% coverage of the theoretical isotopic distribution. Data curation parameters were also applied, including setting the  $m/z$  tolerance was set to  $\pm 100$  ppm and a S/N > 9 was used. Furthermore, the average isotopic pattern quality for each measurement was assessed for the  $[M+2H]^{2+}$ ,  $[M+3H]^{3+}$ , and  $[M+4H]^{4+}$  charge states meeting the above criteria. In this regard, the threshold was set to the average of all charge states plus one s.d. As a result, the isotopic pattern quality threshold was set to a deviation  $\leq 30\%$  from the theoretical isotopic pattern and this cut-off was applied to assess each charge state separately for every analyte. Finally, curated  $[M+2H]^{2+}$ ,  $[M+3H]^{3+}$  and  $[M+4H]^{4+}$  ions were individually quantified, followed by summation to provide a total area of all charge states for each glycan followed by adjustment to 100% in order to cover the complete isotopic envelope.

Only fluorescent peaks which contained at least one *N*-glycan that passed the MS curation parameters were considered for further processing. A third quantification approach, termed as FD-MS, was determined via the crossover of FD and MS data.<sup>16</sup> This was performed by calculating the local relative proportion of *N*-glycan peaks eluting under the same chromatographic peak based on their MS intensities. Then, the fluorescent signal was multiplied by the proportion of each *N*-glycan eluting underneath that chromatographic peak in order to derive an FD-MS signal for each *N*-glycan composition. Furthermore, the FD-MS data for each *N*-glycan composition were used to calculate “derived glycosylation traits”.<sup>14</sup> In this case, direct *N*-glycan traits refers to individual *N*-glycan compositions whereas derived traits describe common structural features shared amongst individual *N*-glycans and are based upon the biosynthetic pathway.<sup>23</sup> These include characteristics such as high-mannose, hybrid-type and complex-type as well as sialylation, galactosylation, and fucosylation. The formula for each calculated derived trait is included in **Supplementary Table S4**. Importantly, the abundances of isomers of the same *N*-glycan composition were summed in order to calculate derived glycosylation traits.

### Statistical Analysis

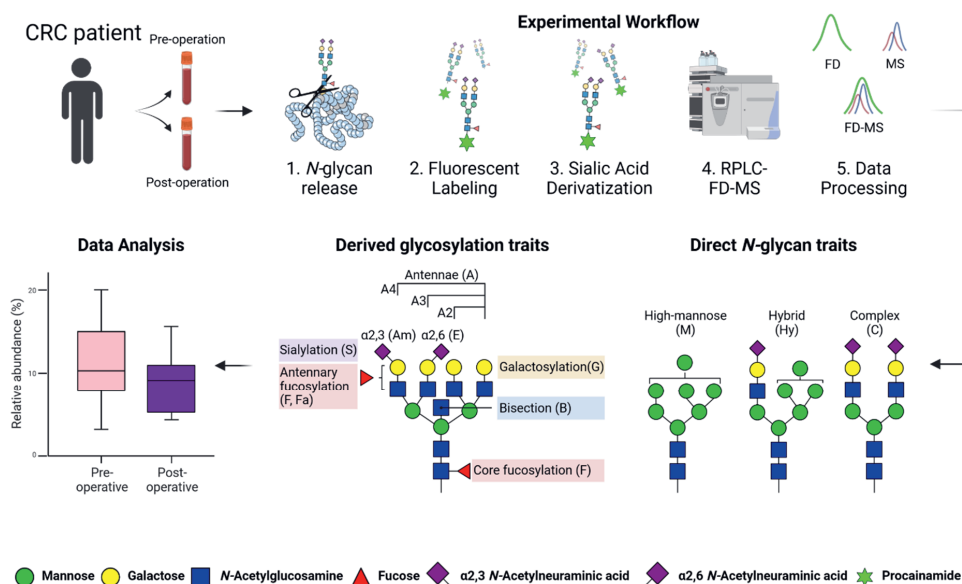
Principal component analysis (PCA) was performed in order to determine the presence of batch effects within the data. Confounders such as age, sex, and sample plate were explored and the batch effect in all three datasets (FD, MS, FD-MS) due to sample plate was corrected using linear regression based on the positive control samples.<sup>24</sup> Direct and derived traits were tested for significance based on the results of the Shapiro-wilk test, a dependent sample T-test or Wilcoxon signed-rank test in order to determine prognostic capability. A significance threshold ( $p$ -value < 0.05) with Bonferroni correction was applied. Univariate analysis was performed in python (Python 3, Create Space, Scotts Valley, CA).

Direct and derived *N*-glycan traits were considered for further analysis when significant differences between pre- and post-operative samples ( $p$ -value < 0.05 with Bonferroni correction) were observed in the current study, as well as being previously validated by the MALDI-MS study.<sup>14</sup> Associations between those specific direct and derived *N*-glycan traits and disease metadata were explored. This included three categories of the disease, namely stage (1 – 4), local extent of tumor assessed pathologically (pT1

– 4), and histological type (Signet-ring cell carcinoma, >50% signet-ring cells; adenocarcinoma, and mucinous adenocarcinoma, >50% mucinous). Thus, to determine whether there was a significant difference of each trait prior to surgery across these categories, a Kruskal-Wallis test was performed with Bonferroni correction, followed by post-hoc analysis looking at the pairwise differences of each subcategory via Dunn test. To gain an understanding of how well a trait is distinguished by a certain category, multiple models were tested and the model with the highest area under the curve (AUC) is reported, while adjusting for age and sex. In this case, the AUC was calculated by comparing one versus rest, in a 5-fold cross validation.

## Results & Discussion

*N*-glycomic profiles of pre- and post-operative serum samples from CRC patients were explored using a newly developed workflow (Figure 1).<sup>16</sup> In this approach, we



**Figure 1. Analysis of pre- and post-operative *N*-glycans in CRC by RPLC-FD-MS.** Serum samples were taken pre- and post-operation from the same patient. Each sample underwent fluorescent labeling and sialic acid derivatization *N*-glycan workflow followed by RPLC-FD-MS measurement. Three quantification approaches were used for data processing, namely FD, MS, and a combination of these two approaches via FD-MS. Direct *N*-glycan and derived glycosylation traits were determined in order to perform data analysis and comparison of the two groups. Adapted with permission from Moran, A.B., *et al.*, 2022. Analytical Chemistry, 94(18), pp.6639-6648 <sup>16</sup>. Copyright 2022 American Chemical Society.

performed fluorescent procainamide labeling and derivatization of  $\alpha$ 2,6- (ethyl esterification) and  $\alpha$ 2,3- (amidation) linked sialic acids followed by RPLC-FD-MS analysis in order to assess serum *N*-glycosylation changes following surgery. In addition to common *N*-glycan features, this method allowed us to perform sialic acid linkage-specific and positional isomer analysis. As shown in **Figure 1**, we investigated direct *N*-glycan traits as well as derived glycosylation traits the latter of which describes global glycosylation characteristics based on the biosynthetic pathway such as sialylation, galactosylation, bisection, fucosylation and antennarity. Unless mentioned otherwise, we refer to FD-MS quantification results throughout this study as we previously demonstrated that this technique achieves a more precise relative quantification and greater coverage of glycan structures than solely performing MS or FD quantification.<sup>16</sup> Furthermore, we compared our results with those of a previous study which analyzed the same cohort of patients using a MALDI-MS approach.<sup>14</sup> In this study, we include a comparison of the same 64 patients analyzed by both methods. Therefore, this allowed us to further validate those results using different ionization, detection and quantification parameters. Finally, we compared the clinical performance of the three quantification techniques that were applied in our study, namely FD, MS and the recently described FD-MS quantification approach.<sup>16</sup>

### **Serum *N*-glycosylation Analysis of Pre- and Post-operative CRC Samples**

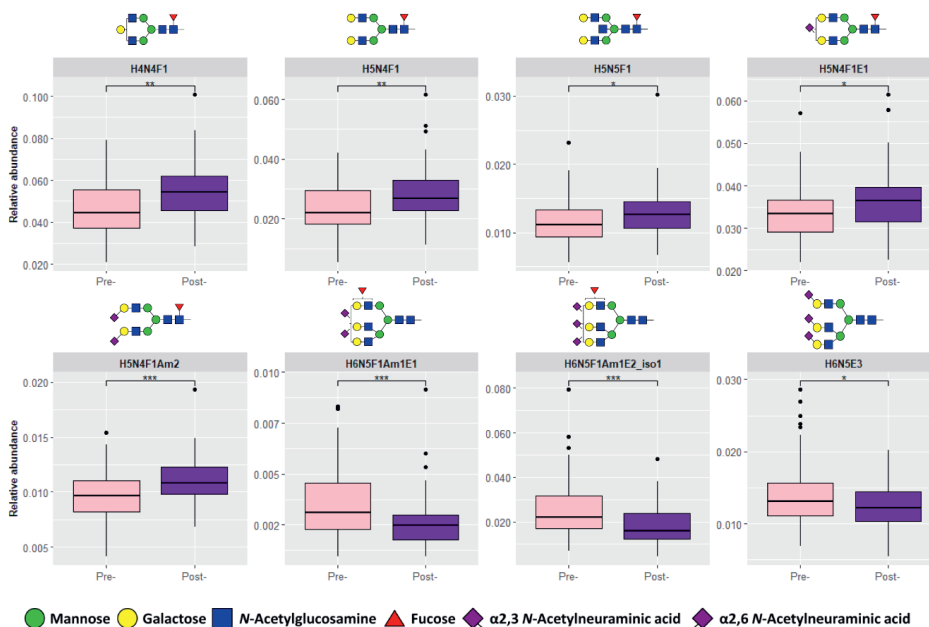
A total of 48 *N*-glycan structures were detected in this study following data curation (**Supplementary Table S5**). Univariate testing was performed in order to assess whether there were any significant differences in the relative abundances of *N*-glycans in pre- and post-operative samples, which resulted in 14 significant *N*-glycans in the FD-MS dataset. **Table 2** shows that 20 direct *N*-glycan traits were previously validated by de Vroome *et al.* using a MALDI-MS approach.<sup>14</sup> Importantly, that study included discovery and validation cohorts using a case/control setup, as well as the same pre- and post-operative CRC cohort that is also assessed by our workflow. From these, 11 *N*-glycans were found in the current study, eight of which were significant discriminators between the pre- and post-operative groups.

**Table 2. Comparison of RPLC-FD-MS and MALDI-MS results.** Asterisk (\*) denotes significant *N*-glycans that were found by the RPLC-FD-MS approach. CRC alteration was calculated as the pre-operative relative abundance divided by the post-operative relative abundance. All results may be found in **Supplementary Table S6**. Abbreviations: T: total, C: Common, n/a: not applicable.

Approach	All <i>N</i> -glycans					Validated <i>N</i> -glycans				
	Count			CRC Alteration		Count		CRC Alteration		
	T	C	Positional isomers	Same	Other	T	C	Significant #	Same	Other
<b>RPLC-FD-MS</b>	48		14	30	9	n/a		8	8	0
<b>MALDI-MS</b>	83	39	0	(77%)	(23%)	20	11	n/a	(100%)	(0%)

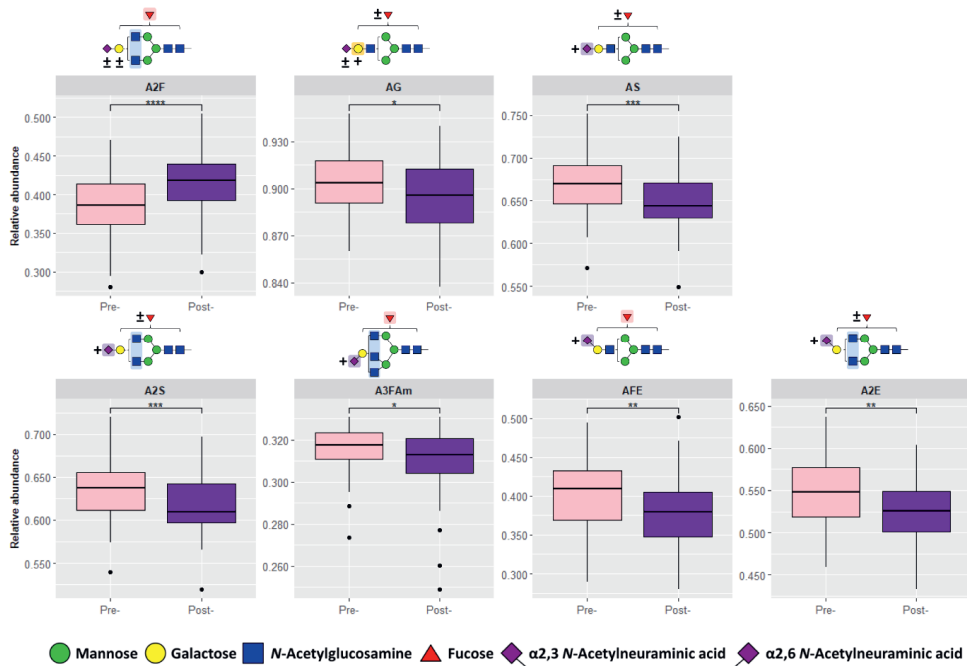
The alteration of these eight significant *N*-glycans in CRC is displayed in **Figure 2**, which shows a lower pre-operative expression of H4N4F1, H5N4F1, H5N4F1S<sub>2,6</sub>1, and H5N4F1S<sub>2,3</sub>2. These results are in line with previous reports that di-antennary structures with core fucosylation are decreased in CRC.<sup>14,25–27</sup> In addition, we also observed that H5N5F1 has a lower abundance in pre-operative CRC, which is supported by previous findings whereby bisected *N*-glycan species were also found to be decreased in serum as well as tissues from CRC patients.<sup>14,25,28</sup> In contrast, a higher abundance of larger structures including H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>1, H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>2 (isomer 1) and H6N5S<sub>2,6</sub>3 in pre-operative CRC is observed. This is in agreement with results which showed that mainly sialylated tri- and tetra-antennary *N*-glycans were found to be increased in CRC.<sup>14,27</sup> Importantly, the same trend in CRC is observed for these structures by the MALDI-MS approach (**Table 2**).

We investigated derived glycosylation traits that describe global features of glycosylation based on the biosynthetic pathway, using the 18 derived traits previously validated by the MALDI-MS study. In this case, we observed 14 common derived traits, of which seven were significant differentiators between pre- and post-operative samples (**Supplementary Table S7**). As shown in **Figure 3**, fucosylated di-antennary glycans (A2F) have a lower expression in pre-operative CRC which is similar to the results of some of the individual direct *N*-glycan traits mentioned above. In addition, α2,6-sialylated fucosylated glycans (AFE) and α2,6-sialylated di-antennary glycans (A2E) have a greater abundance in pre-operative CRC. In relation to this, it has been



**Figure 2. Significant direct N-glycan traits.** Eight significant direct N-glycan traits were found between pre-operative (pre-,  $n = 64$ ) and post-operative (post-,  $n = 64$ ) samples that were also previously validated:<sup>14</sup> H4N4F1, H5N4F1, H5N5F1, H5N4F1S<sub>2,6</sub>1, H5N4F1S<sub>2,3</sub>2, H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>1, H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>2 (isomer 1), and H6N5S<sub>2,6</sub>3. Asterisks denote significance with  $p$ -value < 0.05 (\*), 0.01 (\*\*), 0.001 (\*\*\*). Abbreviations: hexose (H), N-acetylhexosamine (N), fucose (F), amidated  $\alpha$ 2,3-linked N-acetylneuraminic acid (Am) and ethyl esterified  $\alpha$ 2,6-linked N-acetylneuraminic acid (E), isomer (iso).

reported that  $\alpha$ 2,6-sialylation is increased in CRC.<sup>29</sup> Moreover, sialylation per antenna across all glycan species (AS) and sialylated di-antennary N-glycans (A2S) are increased in pre-operative CRC, possibly due the greater abundance of galactosylation per antenna (AG) which would increase the number of available galactose moieties for sialic acid occupation. Furthermore, an increase in  $\alpha$ 2,3-sialylation is observed specifically in fucosylated tri-antennary glycans (A3FAm) in pre-operative CRC. As shown in **Supplementary Table S4**, this derived trait is composed of H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>1 and H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>2 (both isomers summed). **Supplementary Table S1** shows that antennary fucosylation was observed in the case of H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>1 and H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>2 (isomer 1). Thus, this observation supports previous findings that CRC is associated with an increase in sialyl-Lewis epitopes.<sup>30,31</sup> Moreover, Rebello *et al.* analyzed a subset of the same pre- and post-operative cohort



**Figure 3. Significant derived glycosylation traits.** Seven significant derived glycosylation traits were found between pre-operative (pre-,  $n = 64$ ) and post-operative (post-,  $n = 64$ ) samples that were also previously validated:<sup>14</sup> A2F (fucosylated diantennary glycans), AG (galactosylation per antenna), AS (sialic acid per antenna), A2S (sialic acids per diantennary glycan), A3FAm ( $\alpha$ 2,3-linked sialic acids per fucosylated triantennary glycan), AFE ( $\alpha$ 2,6-linked sialic acids per fucosylated glycan), and A2E ( $\alpha$ 2,6-linked sialic acids per diantennary glycan). Asterisks denote significance with  $p$ -value  $< 0.05$  (\*),  $0.01$  (\*\*),  $0.001$  (\*\*\*),  $0.0001$  (\*\*\*\*). Monosaccharide ranges are specified via plus-minus ( $\pm$ ; 0 – 4) and plus (+; 1 – 4) symbols. Abbreviations: antennae (A), fucose (F), galactosylation (G), *N*-acetylneuraminic acid (S), amidated  $\alpha$ 2,3-linked *N*-acetylneuraminic acid (Am) and ethyl esterified  $\alpha$ 2,6-linked *N*-acetylneuraminic acid (E).

using an antennary fucosylation MALDI-MS assay and uncovered that sialyl-Lewis X and Lewis X antigens were significantly increased in pre-operative samples.<sup>32</sup> Thus, our FD-MS results are consistent with those reported by the MALDI-MS<sup>14</sup> approach, indicating the same direction of these seven derived traits in CRC.

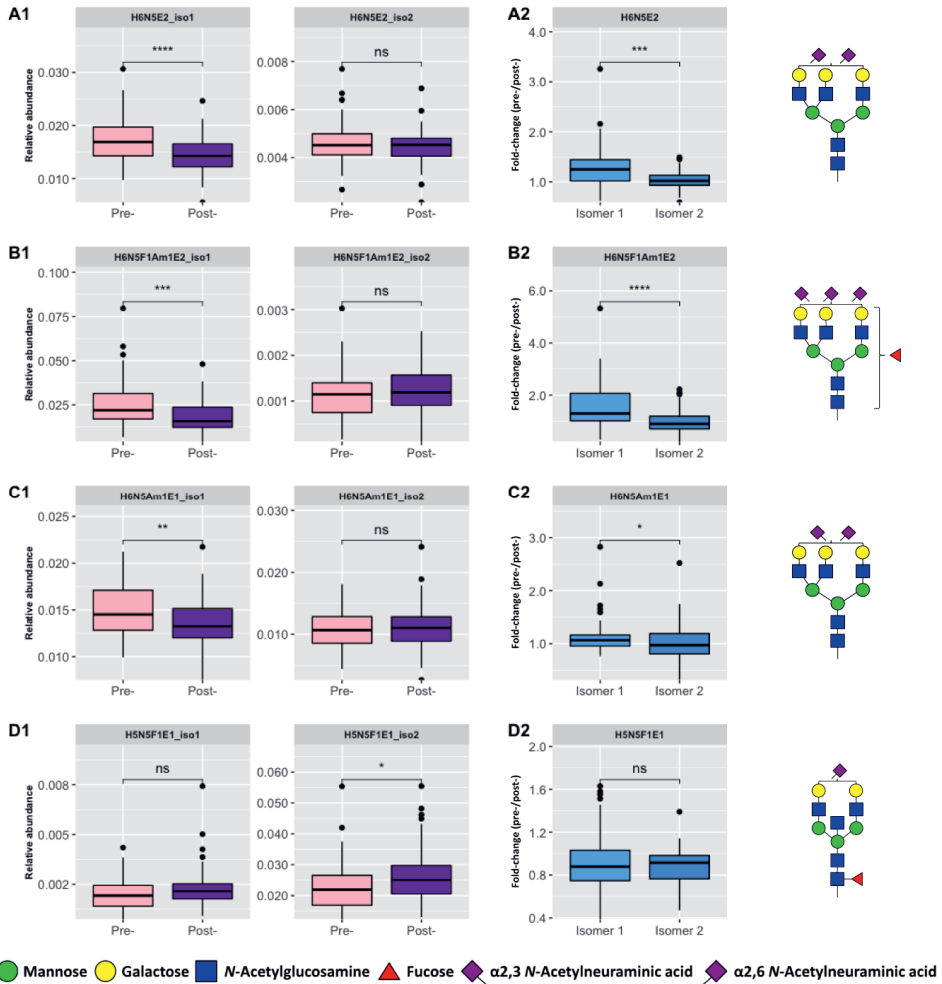
The analysis of the serum *N*-glycome in the pre- and post-operative CRC cohort by MALDI-MS previously resulted in detection of 83 *N*-glycan compositions (Table 2).<sup>14</sup> In comparison, as previously mentioned, we observed a total of 48 *N*-glycan structures in the current study. Whereas an ion trap was utilized in the current workflow it is notable that de Vroome *et al.* employed a Bruker ultrafleXtreme equipped with a TOF mass analyzer operated in reflection mode, which generally demonstrates greater



performance in areas such as mass resolution and mass accuracy.<sup>33</sup> For example, the mass resolution of low, medium and high abundant *N*-glycans measured by the ion trap MS was investigated. In this case, a mass resolution of 6500, 5290 and 6540 was obtained for H4N5 ( $m/z$  951.40), H6N5S<sub>2,3</sub>1S<sub>2,6</sub>2 (isomer 2,  $m/z$  1052.43), and H5N4S<sub>2,6</sub>2 ( $m/z$  834.01), respectively. In relation to this, the isotopic profile of H6N5S<sub>2,3</sub>1S<sub>2,6</sub>2 (isomer 2) is shown in **Supplementary Figure S1**. Takei *et al.* analyzed the serum *N*-glycome using a previous iteration of the Bruker MALDI instrument, the ultraflex III, which resulted in the quantification of 34 *N*-glycans.<sup>25</sup> Thus, it is likely the current workflow could be improved by coupling with a mass spectrometer with higher resolving power and greater sensitivity. Moreover, further improvements in sensitivity could be gained by increasing the injection amount or implementing a nanoflow column.

### Positional Isomers in CRC *N*-glycome

Our workflow demonstrated the efficient chromatographic separation of positional isomers using RPLC following fluorescent labeling and sialic acid derivatization.<sup>16</sup> This is evident in **Table 2** which shows that seven *N*-glycan compositions with the same  $m/z$  were detected as positional isomers, resulting in the assignment of 14 structures in total. We further investigated the structural components which may lead to different retention times for positional isomeric structures. To perform this, we applied our workflow in order to examine IgG *N*-glycans, for which H4N4F1 (G1F,  $m/z$  922.89) is a well-known composition that consists of two isomers.<sup>34</sup> As shown by **Supplementary Figure S2**, we observed a longer retention time for H4N4F1 with  $\alpha$ 3-antenna galactose occupation using a 70 min gradient, indicating that structural variations may lead to different retention times for certain positional isomeric *N*-glycans. Despite this, there is a complex interplay of influences on *N*-glycan retention during RPLC separation, including sialic acid-linkage dependent derivatization, galactose occupation by sialic acids, bisection, number of hexose residues, and fucosylation.<sup>16</sup> As a result, specific assignment of structures remains challenging for *N*-glycans, in particular for those with two antennae or more. Nonetheless, in the case of relevant structures related to a disease, an exoglycosidase approach could be followed in order to characterize specific isomeric features of interest. Following this, this method could then be applied in order to investigate those structures within a clinical context.



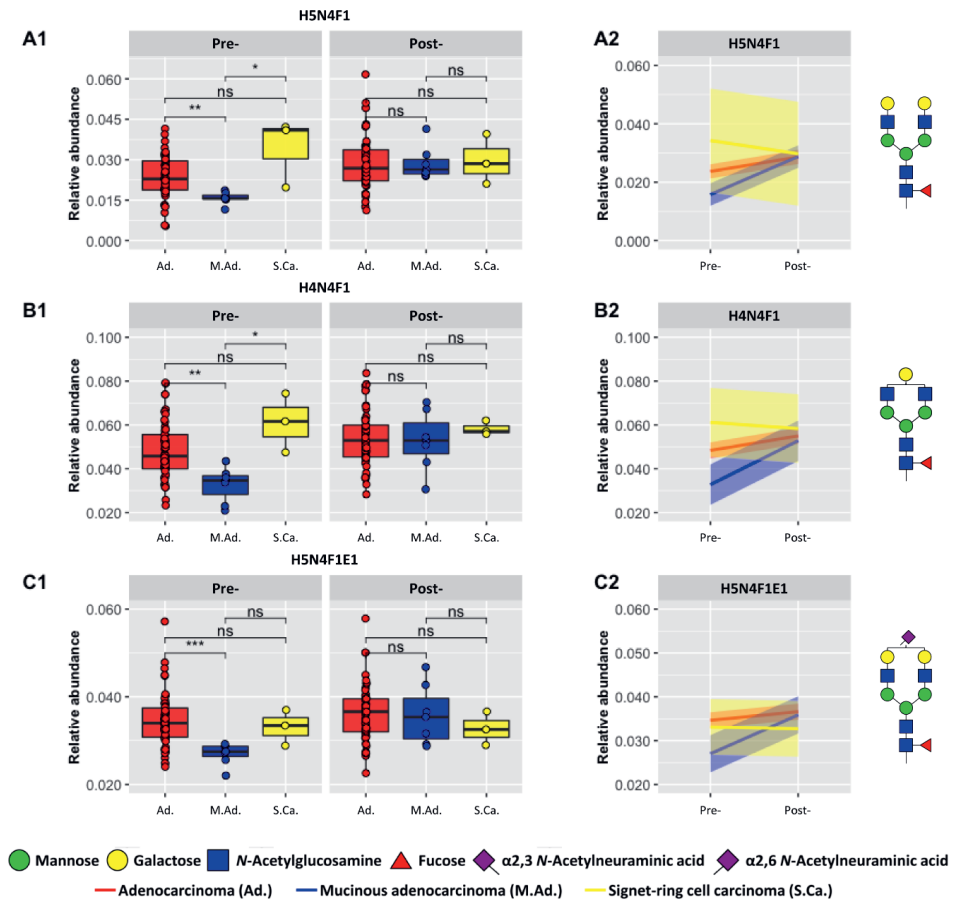
**Figure 4. Positional isomeric N-glycans that show significant differences in CRC.** Four positional isomers with significant differences were observed, as denoted by each row. **(1)** Boxplots of relative abundances observed in pre- ( $n = 64$ ) and post-operative ( $n = 64$ ) samples for the each isomeric form. **(2)** Fold-change in CRC (pre-/post-,  $n = 64$ ) for each isomer whereby an increase ( $>1$ ) or a decrease ( $<1$ ) may be observed. Asterisks denote significance with  $p$ -value  $< 0.05$  (\*),  $0.01$  (\*\*),  $0.001$  (\*\*\*),  $0.0001$  (\*\*\*\*). Abbreviations: Not significant (ns), hexose (H), N-acetylhexosamine (N), fucose (F), amidated  $\alpha$ 2,3-linked N-acetylneuraminic acid (Am) and ethyl esterified  $\alpha$ 2,6-linked N-acetylneuraminic acid (E).

In our study, we revealed that specific isomers are significantly associated with CRC. For example, **Figure 4 (A1 – D1)** showed four compositions where only one of the isomeric forms was a significant differentiator between the pre- and post-operative groups, including a previously validated composition, H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>2. Moreover, **Figure 4 (A2 – D2)** also shows that there were significant differences between the

isomers in CRC fold-change (pre-/post-) for three of these compositions. Previously we showed that antennary fucosylation results in an earlier RT than core fucosylation<sup>16</sup> and, expectedly, **Supplementary Table S5** shows that H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>2 (isomer 1) eluted earlier than H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>2 (isomer 2). In addition, the former assignment is supported by the presence of a fragment ion with  $m/z$  512.20 (**Supplementary Table S1**) which is indicative of antennary fucosylation whereas the latter showed the presence of a core fucose fragment ion ( $m/z$  587.33). In contrast, MALDI-MS may only differentiate mass differences between *N*-glycans. In this case, the location of the fucose residue on H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>2 was not specified by de Vroome *et al.*<sup>14</sup> whereas likely this composition is a mixture of core and antennary fucosylated forms when detected solely based on  $m/z$ . Evidence for fucose positional isomers has been provided by Rebello *et al.* who found a reduction in the relative abundance of H6N5F1S<sub>2,3</sub>1S<sub>2,6</sub>2 following core fucosidase digestion.<sup>32</sup> Thus, here we observed both isomer forms of this *N*-glycan and we illustrated that only the antennary fucosylated form (isomer 1) is significantly different between the pre- and post-operative groups. As a result, our workflow allows the assessment of the specific association of antennary fucosylation with CRC.

### Investigation of Pre- and Post-operative Histological Type

A Kruskal-Wallis test was performed in order to determine whether the pre-operative CRC abundances of the direct and derived *N*-glycan traits shown in **Figures 2 and 3** were significant across three categories of the disease, namely stage, pT and histological type. With regard to this, H5N4F1, H4N4F1 and H5N4F1S<sub>2,6</sub>1 demonstrated significant differences for histological type (**Figure 5**). As shown in **Figure 5 (A1 – C1)**, it can be seen that differences in the pre-operative abundances of these *N*-glycans are no longer observed following surgery (post-operative). In addition, **Figure 5 (A2 – C2)** further demonstrates that the relative abundances converge in the post-operative measurements. Thus, differences in histological type are no longer readily observed following surgery.



**Figure 5. Differences in histological type between pre- and post-operative CRC samples.** Three N-glycans differentiated significantly based upon histological type, adenocarcinoma ( $n = 57$ ), mucinous adenocarcinoma ( $n = 7$ ) and signet-ring cell carcinoma ( $n = 3$ ), as denoted by each row: row A (H5N4F1), row B (H4N4F1) and row C (H5N4F1S<sub>2,6</sub>1). (1) Boxplots of relative abundances observed in pre- and post-operative samples. (2) Trend observed following surgery, from pre- to post-operative samples from the same patients. The 95% CI is shown as colored bands around each line. Asterisks denote significance with  $p$ -value  $< 0.05$  (\*),  $0.01$  (\*\*),  $0.001$  (\*\*\*),  $0.0001$  (\*\*\*\*). Abbreviations: Not significant (ns), hexose (H), N-acetylhexosamine (N), fucose (F), amidated  $\alpha$ 2,3-linked N-acetylneuraminic acid (Am) and ethyl esterified  $\alpha$ 2,6-linked N-acetylneuraminic acid (E).

In order to further support these results, the N-glycans shown in **Figure 2** were investigated retrospectively in the MALDI-MS datasets, namely the pre- ( $n = 64$  pre-operative cases) and post-operative and discovery and validation studies ( $n = 185$  cases). This revealed that the same three N-glycans (H5N4F1, H4N4F1 and H5N4F1S<sub>2,6</sub>1) were also significant differentiators of histological type. As shown in

**Supplementary Figure S3**, a similar trend is observed whereby differences between the subcategories of histological type are no longer observed following surgery. In addition, **Supplementary Figure S4** shows that the results of H5N4F1 and H5N4F1S<sub>2,6</sub>1 were also replicated in cases from the MALDI-MS discovery and validation study. These findings are similar to the results previously demonstrated by de Vroome *et al.* whereby they showed similar *N*-glycan abundances between the post-operative samples and healthy controls,<sup>14</sup> signifying a positive response to surgery.

**Table 3. Comparison of AUCs for adenocarcinoma classification in pre-/post-operative CRC.** AUCs from the comparison of adenocarcinoma ( $n = 57$ ) versus signet-ring cell carcinoma ( $n = 3$ ) and mucinous adenocarcinoma ( $n = 7$ ) are shown. Three quantification approaches performed by this study (RPLC-FD, RPLC-MS, RPLC-FD-MS) are displayed, as well as data from a previous study performing MALDI-MS.<sup>14</sup> "Combination" refers to the combination of the significant features for histological type. All results may be found in **Supplementary Table S8**. Abbreviations: Not detected (n.d.), not applicable (n/a).

Peak	RPLC-FD	Direct <i>N</i> -glycan traits	RPLC-MS	RPLC-FD-MS	MALDI-MS
	AUC		AUC	AUC	AUC
10	0.79	H5N4F1	0.70	0.79	0.79
		H4N4F1	0.72	0.71	0.77
n.d.	n.d.	H5N4F1S <sub>2,6</sub> 1	0.70	0.80	0.76
n.d.	n.d.	Combination	0.68	0.83	0.77

The potential of the *N*-glycans H5N4F1, H4N4F1 and H5N4F1S<sub>2,6</sub>1 as prognostic markers for distinguishing between different histological types was evaluated using a receiver operating characteristic (ROC) curve. As a result of the small sample size, only the largest subcategory versus a combination of the other subcategories is reported. In the case of histological type, **Table 3** displays the classification of adenocarcinoma against signet-ring cell carcinoma and mucinous adenocarcinoma. The analysis of pre- and post-operative CRC samples by the RPLC-FD-MS approach achieved the highest AUC values (ranging from 0.71 to 0.83), and the model combining the three *N*-glycan traits showed the best performance (AUC 0.83). In comparison, the analysis of the same cohort using the MALDI-MS technique demonstrated a moderate performance (AUC 0.76 – 0.79). Although MALDI-MS showed significant differences for pT, its prognostic performance was poor (AUC 0.50 – 0.59; **Supplementary Table S8**). Similarly, the analysis of the MALDI-MS discovery and validation study (**Supplementary Table S8**) revealed significant differences yet poor performance for stratifying adenocarcinoma (AUC 0.56 – 0.60), pT3 (AUC 0.48

– 0.53), and stage 2 (AUC 0.49). The characteristics of the discovery and validation cohort may be found in **Supplementary Table S9**.

Histological type is an important prognostic factor in CRC, as different types of tumors may show different behavior and response to treatment. For example, signet-ring cell carcinoma is associated with poor differentiation (high-grade) and generally results in a worse outcome than adenocarcinomas.<sup>35–38</sup> In addition, microsatellite stable mucinous adenocarcinomas may demonstrate more aggressive behavior when detected at an advanced stage.<sup>35,39</sup> Undoubtedly, identifying the histological type of CRC can assist in guiding the treatment approach as well as predicting patient prognosis<sup>35</sup> and yet, interobserver variability can present challenges in this regard.<sup>40</sup> Thus, the results displayed by the current study demonstrate the utility of the serum *N*-glycome in aiding histological type classification. Moreover, we also illustrated that the response to surgery may be monitored by comparing pre- and post-operative *N*-glycan abundances from the same patient, thus promoting an individualized approach to treatment. This is an important result as the need and benefit of personalized medicine is continuously being recognized.<sup>41,42</sup> Despite this, it should be noted that this cohort consisted mainly of the most prominent histological type, adenocarcinoma<sup>43</sup> (**Table 1**), and as a result, it may be challenging to translate these results to subsequent studies. Nevertheless, this analysis shows the potential of the serum *N*-glycome as a prognostic marker and further validation should be performed, particularly with regard to histological type classification and response to surgery.

### Comparison of Clinical Performance of Three Quantification Approaches

In this study we employed three quantification approaches, FD, MS, and FD-MS, in order to analyze pre- and post-operative CRC. Previously we showed that FD-MS quantification, due to the detection of specific glycan masses, resulted in a greater coverage of glycan features than FD. Furthermore, as a result of incorporating the FD signal into the quantification of individual *N*-glycans, we also demonstrated that FD-MS achieved greater precision than MS quantification.<sup>16</sup> In the current work, we further compared the performance of these quantification approaches in a clinical setting. In **Supplementary Table S5**, the number of significant peaks or glycan structures is provided for the FD (14), MS (11), and FD-MS (14) approaches. In relation to this, the number of validated significant *N*-glycans for MS and FD-MS is seven and eight,

respectively. Furthermore, there were seven FD peaks (10, 11, 13, 17, 24, 25, 32) that were significant between the groups and also contained a previously validated *N*-glycan. Thus, it is shown here that the three quantification approaches perform similarly in terms of number of validated and significant features (FD peak or glycan structure). In addition, the three quantification approaches were also evaluated for their diagnostic performance over time in relation to adenocarcinoma classification via ROC curves whereby AUCs were also calculated (**Table 3**). This demonstrates a good performance for FD-MS and FD quantification which both fall into a similar AUC range (AUC 0.70 – 0.80). In comparison, MS quantification resulted in a fair performance (0.68 – 0.72). Thus, in comparison with FD, the FD-MS quantification approach showed a similar performance yet allowed the analysis of specific direct *N*-glycan traits in relation to CRC. The comparison of MS and FD-MS demonstrated that the latter showed improved diagnostic performance for adenocarcinoma classification.

Cancer-associated alterations in glycosylation have been well established in the literature.<sup>44–46</sup> Moreover, several studies have demonstrated the potential with regard to the application of glycomics for aiding and improving cancer diagnosis and prognosis.<sup>14,47,48</sup> Despite this, there is still a lack of glycomic tests that have been translated from biomedical research into clinical laboratories.<sup>49</sup> This is partly because of the challenges that carbohydrates themselves present for analytical method development, detection and quantification, as well as a lack of validation studies beyond the discovery phase.<sup>50</sup> In this study, we have demonstrated that MS-detection allows the underlying glycosylation alterations in CRC to be probed. As a result, a greater understanding of the disease may be developed whilst maintaining specificity with regard to the traits that can be targeted for diagnostic and prognostic testing. However, FD of labeled glycans shows promise for clinical development<sup>51</sup> as it may avoid common challenges with glycan analysis by MS, such as poor ionization. In this regard, we have demonstrated similar and satisfactory performance of FD and FD-MS quantification (**Table 3**). Moreover, FD-MS relative abundances more closely resemble FD relative abundances in comparison with MS-based quantification (**Supplementary Figure S5**). Thus, upon identification of the best-performing *N*-glycan traits, the FD-MS results could be more easily bridged with the FD platform, enabling efficient implementation within the clinical setting.

## Perspectives

Undoubtedly, as previously mentioned, the results presented here require further validation with a larger and more diverse cohort of CRC patients in order to evaluate the true prognostic potential of serum *N*-glycosylation. In this regard, the monitoring of patient profiles before, during and after clinical events such as treatment, therapy or surgery seems promising and should be further explored. In addition, clinical validation of FD-MS quantification and bridging to FD quantification should also be examined. In this regard, interferences that may affect MS and FD signals, such as different mobile phase conditions across the gradient as well as unidentified fluorescent compounds, should be defined and, if possible, eliminated. Overall, the presented workflow is suitable for future investigations into biomarker discovery as well as clinical validation and translation.

## Conclusions

In this study, we evaluated the clinical performance of our workflow which performs *N*-glycan fluorescent labeling and sialic acid derivatization followed by RPLC-FD-MS measurement. Under alternative ionization, detection and quantification conditions, we further supported previously validated direct *N*-glycan and derived glycosylation traits for differentiating pre-operative CRC in comparison with post-operative samples from the same patient. Furthermore, we observed that specific positional isomers were significant between the two groups which has important clinical implications in the context of this disease. In addition, pre-operative abundances of *N*-glycans showed good performance for the classification of adenocarcinoma and new clinical insights were gained as we illustrated that differences between histological types were abolished following surgery. Thus, serum *N*-glycosylation is a promising CRC prognostic marker that may be further investigated for performing disease classification and monitoring patient response to surgery.

## Supplementary Materials

**Supplementary figures (.pdf):** **Figure S1.** Representative profile of *N*-glycans measured by RPLC-FD-MS. **Figure S2.** Extracted ion chromatograms of H4N4F1 from IgG. **Figure S3.** Differences in histological type between pre- and post-operative



samples measured by MALDI-MS. **Figure S4.** Significant *N*-glycans for histological type in MALDI-MS discovery and validation study. **Figure S5.** Linear regression comparison.

**Supplementary tables (xlsx):** **Table S1.** Diagnostic ions obtained by RPLC-FD-MS(-MS). **Table S2.** HappyTools peak extraction and calibration parameters. **Table S3.** LaCyTools composition extraction and calibration parameters. **Table S4.** Formulas for calculating derived glycosylation traits. **Table S5.** Analysis of CRC cohort using three quantification methods. **Table S6.** Comparison of RPLC-FD-MS and MALDI-MS direct *N*-glycan trait results. **Table S7.** Comparison of RPLC-FD-MS and MALDI-MS derived glycosylation trait results. **Table S8.** Diagnostic performance of RPLC- (FD, MS, FD-MS) and MALDI-MS results for discriminating CRC disease categories. **Table S9.** Clinical characteristics of the discovery and validation CRC cohort.

## Author Contributions

Conceptualization, D.I.R.S., M.W., Y.E.M.v d B., G.S.M.L.K. and A.B.M.; methodology, A.B.M.; data curation, A.B.M.; formal analysis, A.B.M., G.E.H., Y.E.M.v d B., and G.S.M.L.K.; resources, W.E.M., R.A.E.M.T., D.I.R.S. and M.W.; writing—original draft preparation, A.B.M., G.S.M.L.K., G.E.H., Y.E.M.v d B., and M.W.; writing—review and editing, A.B.M., G.S.M.L.K., G.E.H., Y.E.M.v d B., M.W., W.E.M., R.A.E.M.T., and D.I.R.S.; visualization, A.B.M., and G.S.M.L.K.; supervision, G.S.M.L.K., Y.E.M.v d B., M.W., and D.I.R.S.; funding acquisition, G.S.M.L.K., D.I.R.S., and M.W. All authors have read and agreed to the published version of the manuscript.

## Funding

This research was funded by the European Union's Horizon 2020 Research and Innovation Program (GlySign, Grant No. 722095).

## Institutional Review Board Statement

The study was conducted in accordance with the Declaration of Helsinki. Serum samples were collected at the Leiden University Medical Center (LUMC) Surgical Oncology Biobank between October 2002 and March 2013 according to a

standardized protocol. This study was approved by the Medical Ethics Committee of the LUMC and was performed in agreement with the Dutch law for medical research involving human subjects (see also Code of Conduct of the Federation of Medical Scientific Societies in the Netherlands <http://www.federa.org/>).

## Informed Consent Statement

Informed consent was obtained from all subjects involved in the study.

## Data Availability Statement

The raw data is available from the authors upon request.

## Acknowledgments

The authors wish to acknowledge that Figure 1 was created using Biorender.com.

## Conflicts of Interest

The authors declare the following competing financial interest(s): Daniel I. R. Spencer and Georgia Elgood-Hunt are employees of Ludger Ltd., a company that provides commercial glycoanalytical products and services.

## References

- (1) Vajaria, B. N.; Patel, P. S. Glycosylation: A Hallmark of Cancer? *Glycoconjugate Journal*. **2017**, *34* (2). DOI: 10.1007/s10719-016-9755-2.
- (2) Silsirivanit, A. Glycosylation Markers in Cancer. In *Advances in Clinical Chemistry*; **2019**; Vol. 89. DOI: 10.1016/bs.acc.2018.12.005.
- (3) Ohtsubo, K.; Marth, J. D. Glycosylation in Cellular Mechanisms of Health and Disease. *Cell*. **2006**, *126* (5). DOI: 10.1016/j.cell.2006.08.019.
- (4) Munkley, J.; Elliott, D. J. Hallmarks of Glycosylation in Cancer. *Oncotarget*. **2016**, *7* (23). DOI: 10.18632/oncotarget.8155.
- (5) Pinho, S. S.; Reis, C. A. Glycosylation in Cancer: Mechanisms and Clinical Implications. *Nature Reviews Cancer*. **2015**, *15* (9). DOI: 10.1038/nrc3982.
- (6) Boyaval, F.; Dalebout, H.; Van Zeijl, R.; Wang, W.; Fariña-Sarasqueta, A.; Lageveen-Kammeijer, G. S. M.; Boonstra, J. J.; McDonnell, L. A.; Wuhler, M.; Morreau, H.; Heijs, B. High-Mannose N-Glycans as Malignant Progression Markers in Early-Stage Colorectal Cancer. *Cancers*. **2022**, *14* (6). DOI: 10.3390/cancers14061552.

- (7) Boyaval, F.; Van Zeijl, R.; Dalebout, H.; Holst, S.; Van Pelt, G.; Fariña-Sarasqueta, A.; Mesker, W.; Tollenaar, R.; Morreau, H.; Wuhrer, M.; Heijs, B. N-Glycomic Signature of Stage II Colorectal Cancer and Its Association with the Tumor Microenvironment. *Molecular and Cellular Proteomics*. **2021**, *20*. DOI: 10.1074/MCP.RA120.002215.
- (8) Drake, R. R.; Powers, T. W.; Jones, E. E.; Bruner, E.; Mehta, A. S.; Angel, P. M. MALDI Mass Spectrometry Imaging of N-Linked Glycans in Cancer Tissues. In *Advances in Cancer Research*; **2017**; Vol. 134. DOI: 10.1016/bs.acr.2016.11.009.
- (9) Houvast, R. D.; Vankemmelbeke, M.; Durrant, L. G.; Wuhrer, M.; Baart, V. M.; Kuppen, P. J. K.; de Geus-Oei, L. F.; Vahrmeijer, A. L.; Sier, C. F. M. Targeting Glycans and Heavily Glycosylated Proteins for Tumor Imaging. *Cancers*. **2020**, *12* (12). DOI: 10.3390/cancers12123870.
- (10) Holm, M.; Nummela, P.; Heiskanen, A.; Satomaa, T.; Kaprio, T.; Mustonen, H.; Ristimäki, A.; Haglund, C. N-Glycomic Profiling of Colorectal Cancer According to Tumor Stage and Location. *PLoS ONE*. **2020**, *15* (6 June). DOI: 10.1371/journal.pone.0234989.
- (11) Pan, Y.; Zhang, L.; Zhang, R.; Han, J.; Qin, W.; Gu, Y.; Sha, J.; Xu, X.; Feng, Y.; Ren, Z.; Dai, J.; Huang, B.; Ren, S.; Gu, J. Screening and Diagnosis of Colorectal Cancer and Advanced Adenoma by Bionic Glycome Method and Machine Learning. *Am J Cancer Res*. **2021**, *11* (6).
- (12) Bertok, T.; Bertokova, A.; Jane, E.; Hires, M.; Aguedo, J.; Potocarova, M.; Lukac, L.; Vikartovska, A.; Kasak, P.; Borsig, L.; Tkac, J. Identification of Whole-Serum Glycobiomarkers for Colorectal Carcinoma Using Reverse-Phase Lectin Microarray. *Frontiers in Oncology*. **2021**, *11*. DOI: 10.3389/fonc.2021.735338.
- (13) Crosby, D.; Lyons, N.; Greenwood, E.; Harrison, S.; Hiom, S.; Moffat, J.; Quallo, T.; Samuel, E.; Walker, I. A Roadmap for the Early Detection and Diagnosis of Cancer. *The Lancet Oncology*. **2020**, *21* (11). DOI: 10.1016/S1470-2045(20)30593-3.
- (14) de Vroome, S. W.; Holst, S.; Gironde, M. R.; van der Burgt, Y. E. M.; Mesker, W. E.; Tollenaar, R. A. E. M.; Wuhrer, M. Serum N-Glycome Alterations in Colorectal Cancer Associate with Survival. *Oncotarget*. **2018**, *9* (55), 30610–30623. DOI: 10.18632/oncotarget.25753.
- (15) Vreeker, G. C. M.; Nicolardi, S.; Bladergroen, M. R.; Van Der Plas, C. J.; Mesker, W. E.; Tollenaar, R. A. E. M.; Van Der Burgt, Y. E. M.; Wuhrer, M. Automated Plasma Glycomics with Linkage-Specific Sialic Acid Esterification and Ultrahigh Resolution MS. *Analytical Chemistry*. **2018**, *90* (20). DOI: 10.1021/acs.analchem.8b02391.
- (16) Moran, A. B.; Gardner, R. A.; Wuhrer, M.; Lageveen-Kammeijer, G. S. M.; Spencer, D. I. R. Sialic Acid Derivatization of Fluorescently Labeled N-Glycans Allows Linkage Differentiation by Reversed-Phase Liquid Chromatography–Fluorescence Detection–Mass Spectrometry. *Analytical Chemistry*. **2022**, *94* (18), 6639–6648. DOI: 10.1021/acs.analchem.1c02610.
- (17) Reiding, K. R.; Blank, D.; Kuijper, D. M.; Deelder, A. M.; Wuhrer, M. High-Throughput Profiling of Protein N-Glycosylation by MALDI-TOF-MS Employing Linkage-Specific Sialic Acid Esterification. *Analytical Chemistry*. **2014**, *86* (12), 5784–5793. DOI: 10.1021/ac500335t.
- (18) Lageveen-Kammeijer, G. S. M.; de Haan, N.; Mohaupt, P.; Wagt, S.; Filius, M.; Nouta, J.; Falck, D.; Wuhrer, M. Highly Sensitive CE-ESI-MS Analysis of N-Glycans from Complex

- Biological Samples. *Nature Communications*. **2019**, *10* (1), 1–8. DOI: 10.1038/s41467-019-09910-7.
- (19) Varki, A. et al. Symbol Nomenclature for Graphical Representations of Glycans. *Glycobiology*. **2015**, *25* (12), 1323–1324. DOI: 10.1093/glycob/cwv091.
- (20) Jansen, B. C.; Hafkenscheid, L.; Bondt, A.; Gardner, R. A.; Hendel, J. L.; Wuhrer, M.; Spencer, D. I. R. HappyTools: A Software for High-Throughput HPLC Data Processing and Quantitation. *PLoS One*. **2018**, *13* (7), e0200280. DOI: 10.1371/journal.pone.0200280.
- (21) European Medicines Agency (EMA). *ICH Guidelines Q2(R2) on Validation of Analytical Procedures*; **2022**; Vol. 2.
- (22) Jansen, B. C.; Falck, D.; de Haan, N.; Hipgrave Ederveen, A. L.; Razdorov, G.; Lauc, G.; Wuhrer, M. LaCyTools: A Targeted Liquid Chromatography-Mass Spectrometry Data Processing Package for Relative Quantitation of Glycopeptides. *Journal of Proteome Research*. **2016**, *15* (7). DOI: 10.1021/acs.jproteome.6b00171.
- (23) Colley KJ, Varki A, K. T. Cellular Organization of Glycosylation In: *Essentials of Glycobiology*. **2017**.
- (24) Wehrens, R.; Hageman, J. A.; van Eeuwijk, F.; Kooke, R.; Flood, P. J.; Wijnker, E.; Keurentjes, J. J. B.; Lommen, A.; van Eekelen, H. D. L. M.; Hall, R. D.; Mumm, R.; de Vos, R. C. H. Improved Batch Correction in Untargeted MS-Based Metabolomics. *Metabolomics*. **2016**, *12* (5). DOI: 10.1007/s11306-016-1015-8.
- (25) Takei, D.; Harada, K.; Nouse, K.; Miyahara, K.; Dohi, C.; Matsushita, H.; Kinugasa, H.; Hiraoka, S.; Nishimura, S. I.; Okada, H. Clinical Utility of a Serum Glycome Analysis in Patients with Colorectal Cancer. *Journal of Gastroenterology and Hepatology (Australia)*. **2022**, *37* (4). DOI: 10.1111/jgh.15781.
- (26) Zhao, Y. P.; Ruan, C. P.; Wang, H.; Hu, Z. Q.; Fang, M.; Gu, X.; Ji, J.; Zhao, J. Y.; Gao, C. F. Identification and Assessment of New Biomarkers for Colorectal Cancer with Serum N-Glycan Profiling. *Cancer*. **2012**, *118* (3). DOI: 10.1002/cncr.26342.
- (27) Doherty, M.; Theodoratou, E.; Walsh, I.; Adamczyk, B.; Stöckmann, H.; Agakov, F.; Timofeeva, M.; Trbojević-Akmačić, I.; Vučković, F.; Duffy, F.; McManus, C. A.; Farrington, S. M.; Dunlop, M. G.; Perola, M.; Lauc, G.; Campbell, H.; Rudd, P. M. Plasma N-Glycans in Colorectal Cancer Risk. *Scientific Reports*. **2018**, *8* (1). DOI: 10.1038/s41598-018-26805-7.
- (28) Balog, C. I. A.; Stavenhagen, K.; Fung, W. L. J.; Koeleman, C. A.; McDonnell, L. A.; Verhoeven, A.; Mesker, W. E.; Tollenaar, R. A. E. M.; Deelder, A. M.; Wuhrer, M. N-Glycosylation of Colorectal Cancer Tissues. *Molecular & Cellular Proteomics*. **2012**, *11* (9). DOI: 10.1074/mcp.m111.011601.
- (29) Park, J. J.; Lee, M. Increasing the  $\alpha$  2, 6 Sialylation of Glycoproteins May Contribute to Metastatic Spread and Therapeutic Resistance in Colorectal Cancer. *Gut and Liver*. **2013**, *7* (6). DOI: 10.5009/gnl.2013.7.6.629.
- (30) Nakagoe, T.; Sawai, T.; Tsuji, T.; Jibiki, M. A.; Nanashima, A.; Yamaguchi, H.; Yasutake, T.; Kurosaki, N.; Ayabe, H.; Arisawa, K. Preoperative Serum Levels of Sialyl Lewis(a), Sialyl Lewis(x), and Sialyl Tn Antigens as Prognostic Markers after Curative Resection for Colorectal Cancer. *Cancer detection and prevention*. **2001**, *25* (3).

- (31) Miyoshi, E.; Moriwaki, K.; Nakagawa, T. Biological Function of Fucosylation in Cancer Biology. *Journal of Biochemistry*. **2008**, *143* (6). DOI: 10.1093/jb/mvn011.
- (32) Rebello, O. D.; Nicolardi, S.; Lageveen-Kammeijer, G. S. M.; Nouta, J.; Gardner, R. A.; Mesker, W. E.; Tollenaar, R. A. E. M.; Spencer, D. I. R.; Wuhrer, M.; Falck, D. A Matrix-Assisted Laser Desorption/Ionization—Mass Spectrometry Assay for the Relative Quantitation of Antennary Fucosylated N-Glycans in Human Plasma. *Frontiers in Chemistry*. **2020**, *8*. DOI: 10.3389/fchem.2020.00138.
- (33) de Hoffmann, E.; Stroobant, V. *Mass Spectrometry: Principles and Applications*, 3rd ed.; Wiley: Chichester, **2007**.
- (34) Keser, T.; Pavić, T.; Lauc, G.; Gornik, O. Comparison of 2-Aminobenzamide, Procainamide and RapiFluor-MS as Derivatizing Agents for High-Throughput HILIC-UPLC-FLR-MS N-Glycan Analysis. *Frontiers in Chemistry*. **2018**, *6*. DOI: 10.3389/fchem.2018.00324.
- (35) Fleming, M.; Ravula, S.; Tatishchev, S. F.; Wang, H. L. Colorectal Carcinoma: Pathologic Aspects. *Journal of Gastrointestinal Oncology*. **2012**, *3* (3). DOI: 10.3978/j.issn.2078-6891.2012.030.
- (36) Makino, T.; Tsujinaka, T.; Mishima, H.; Ikenaga, M.; Sawamura, T.; Nakamori, S.; Fujitani, K.; Hirao, M.; Kashiwazaki, M.; Masuda, N.; Takeda, M.; Mano, M. Primary Signet-Ring Cell Carcinoma of the Colon and Rectum: Report of Eight Cases and Review of 154 Japanese Cases. *Hepato-Gastroenterology*. **2006**, *53* (72).
- (37) Chen, J. S.; Hsieh, P. S.; Chiang, J. M.; Yeh, C. Y.; Tsai, W. S.; Tang, R.; Changchien, C. R.; Wu, R. C. Clinical Outcome of Signet Ring Cell Carcinoma and Mucinous Adenocarcinoma of the Colon. *Chang Gung Medical Journal*. **2010**, *33* (1).
- (38) Kang, H.; O'Connell, J. B.; Maggard, M. A.; Sack, J.; Ko, C. Y. A 10-Year Outcomes Evaluation of Mucinous and Signet-Ring Cell Carcinoma of the Colon and Rectum. *Diseases of the Colon and Rectum*. **2005**, *48* (6). DOI: 10.1007/s10350-004-0932-1.
- (39) Leopoldo, S.; Lorena, B.; Cinzia, A.; Gabriella, D. C.; Angela Luciana, B.; Renato, C.; Antonio, M.; Carlo, S.; Cristina, P.; Stefano, C.; Maurizio, T.; Luigi, R.; Cesare, B. Two Subtypes of Mucinous Adenocarcinoma of the Colorectum: Clinicopathological and Genetic Features. *Annals of Surgical Oncology*. **2008**, *15* (5). DOI: 10.1245/s10434-007-9757-1.
- (40) Compton, C. C.; Fielding, L. P.; Burgart, L. J.; Conley, B.; Cooper, H. S.; Hamilton, S. R.; Hammond, M. E. H.; Henson, D. E.; Hutter, R. V. P.; Nagle, R. B.; Nielsen, M. L.; Sargent, D. J.; Taylor, C. R.; Welton, M.; Willett, C. Prognostic Factors in Colorectal Cancer: College of American Pathologists Consensus Statement 1999. In *Archives of Pathology and Laboratory Medicine*; **2000**; Vol. 124.
- (41) Hanna-Sawires, R. G.; Schiphuis, J. H.; Wuhrer, M.; Vasen, H. F. A.; van Leerdam, M. E.; Bonsing, B. A.; Mesker, W. E.; van der Burgt, Y. E. M.; Tollenaar, R. A. E. M. Clinical Perspective on Proteomic and Glycomic Biomarkers for Diagnosis, Prognosis, and Prediction of Pancreatic Cancer. *International Journal of Molecular Sciences*. **2021**, *22* (5). DOI: 10.3390/ijms22052655.
- (42) Jackson, S. E.; Chester, J. D. Personalised Cancer Medicine. *International Journal of Cancer*. **2015**, *137* (2). DOI: 10.1002/ijc.28940.

- (43) Stewart, S. L.; Wike, J. M.; Kato, I.; Lewis, D. R.; Michaud, F. A Population-Based Study of Colorectal Cancer Histology in the United States, 1998-2001. *Cancer*. **2006**, *107* (SUPPL.). DOI: 10.1002/cncr.22010.
- (44) Mereiter, S.; Balmaña, M.; Campos, D.; Gomes, J.; Reis, C. A. Glycosylation in the Era of Cancer-Targeted Therapy: Where Are We Heading? *Cancer Cell*. **2019**, *36* (1). DOI: 10.1016/j.ccell.2019.06.006.
- (45) Lau, K. S.; Partridge, E. A.; Grigorian, A.; Silvescu, C. I.; Reinhold, V. N.; Demetriou, M.; Dennis, J. W. Complex N-Glycan Number and Degree of Branching Cooperate to Regulate Cell Proliferation and Differentiation. *Cell*. **2007**, *129* (1). DOI: 10.1016/j.cell.2007.01.049.
- (46) Hang, Q.; Isaji, T.; Hou, S.; Zhou, Y.; Fukuda, T.; Gu, J. N-Glycosylation of Integrin A5 Acts as a Switch for EGFR-Mediated Complex Formation of Integrin A5 $\beta$ 1 to A6 $\beta$ 4. *Scientific Reports*. **2016**, *6*. DOI: 10.1038/srep33507.
- (47) Yoneyama, T.; Ohyama, C.; Hatakeyama, S.; Narita, S.; Habuchi, T.; Koie, T.; Mori, K.; Hidari, K. I. P. J.; Yamaguchi, M.; Suzuki, T.; Tobisawa, Y. Measurement of Aberrant Glycosylation of Prostate Specific Antigen Can Improve Specificity in Early Detection of Prostate Cancer. *Biochemical and Biophysical Research Communications*. **2014**. DOI: 10.1016/j.bbrc.2014.04.107.
- (48) Vreeker, G. C. M.; Hanna-Sawires, R. G.; Mohammed, Y.; Bladergroen, M. R.; Nicolardi, S.; Dotz, V.; Nouta, J.; Bonsing, B. A.; Mesker, W. E.; van der Burgt, Y. E. M.; Wuhrer, M.; Tollenaar, R. A. E. M. Serum N-Glycome Analysis Reveals Pancreatic Cancer Disease Signatures. *Cancer Medicine*. **2020**, *9* (22). DOI: 10.1002/cam4.3439.
- (49) Almeida, A.; Kolarich, D. The Promise of Protein Glycosylation for Personalised Medicine. *Biochimica et Biophysica Acta - General Subjects*. **2016**, *1860* (8). DOI: 10.1016/j.bbagen.2016.03.012.
- (50) de Haan, N.; Wuhrer, M.; Ruhaak, L. R. Mass Spectrometry in Clinical Glycomics: The Path from Biomarker Identification to Clinical Implementation. *Clinical Mass Spectrometry*. **2020**, *18*. DOI: 10.1016/j.clinms.2020.08.001.
- (51) Colhoun, H. O.; Treacy, E. P.; MacMahon, M.; Rudd, P. M.; Fitzgibbon, M.; O'Flaherty, R.; Stepien, K. M. Validation of an Automated Ultraperformance Liquid Chromatography IgG N-Glycan Analytical Method Applicable to Classical Galactosaemia. *Annals of Clinical Biochemistry*. **2018**, *55* (5). DOI: 10.1177/0004563218762957.

