



OPEN Combination of urinary free-glycan markers for the diagnosis of various malignant tumors

Miki Tanaka-Okamoto¹, Ken Hanzawa², Takashi Yamamoto³, Tomoki Michida⁴, Kenji Ikezawa⁵, Kazuyoshi Ohkawa⁵, Kazumi Nishino⁶, Takafumi Yokota⁷, Michihide Maeda⁸, Ken-ichi Yoshida⁹, Satoshi Takenaka¹⁰, Masayuki Ohue¹¹, Tomoyuki Yamasaki³ & Yasuhide Miyamoto³✉

Many urinary free-glycan markers were identified in our previous studies. Here, the clinical utility of these markers was examined. Urine samples taken from 120 healthy subjects and 503 patients with various malignant tumors were analyzed. Four lactose-core glycans containing *N*-glycolylglucosamine (GlcNGc) were synthesized and used as internal standards (ISs). Free-glycans were isolated using a two-step column purification procedure, pyridylaminated, and subjected to LC-selected reaction monitoring. Assay validation was performed using four ISs and eight reference glycans. Twelve markers composed of three sialyl lactose-core glycans, three sulfated glycans, four Gn1-core free-*N*-glycans, and two glycans with unusual structures were selected to investigate their effectiveness for cancer diagnosis. Markers were simultaneously measured and the relative area ratio (marker/IS) quantified. This method was shown to be accurate and precise by repeated analysis of calibrators and quality control samples in urine. Receiver operating characteristic curve analyses revealed that individual markers were not sufficient for highly accurate diagnosis. However, a combination of four markers resulted in higher area under curves of 0.910, 0.867, and 0.914 for gastric, colorectal, and pancreatic cancers, respectively. Moreover, levels of these markers were elevated in various other malignancies. These analyses demonstrated high clinical utility of the free-glycan markers.

Keywords Cancer, Free-glycan, Marker, SRM, Urine

Abbreviations

LC	Liquid chromatography
SRM	Selected reaction monitoring
IS	Internal standard
GlcNGc	<i>N</i> -Glycolylglucosamine
HPLC	High performance liquid chromatography
ROC	Receiver operating characteristic
AUC	Area under curve
GlcNAc	<i>N</i> -Acetylglucosamine
SLA	Sialyl Lewis A tetra

¹Department of Glyco-Oncology and Medical Biochemistry, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ²Glycoanalytical Chemistry Lab, Institute for Glyco-Core Research (iGCORE), Gifu University, 1-1 Yanagido, Gifu 501-1193, Japan. ³Department of Clinical Laboratory, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ⁴Department of Gastrointestinal Oncology, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ⁵Department of Hepatobiliary and Pancreatic Oncology, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ⁶Department of Thoracic Oncology, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ⁷Department of Hematology, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ⁸Department of Gynecology, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ⁹Department of Diagnostic Pathology and Cytology, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ¹⁰Department of Orthopedic Surgery, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ¹¹Department of Gastroenterological Surgery, Osaka International Cancer Institute, 3-1-69 Otemae, Chuo-Ku, Osaka 541-8567, Japan. ✉email: yasuhide.miyamoto@oici.jp

LST-b	LS-tetrasaccharide b
LST-c	LS-tetrasaccharide c
LNnH	Lacto- <i>N</i> -neohexaose
DFLNH(a)	Difucosyl-lacto- <i>N</i> -hexaose (a)
DFLNnH	Difucosyl-lacto- <i>N</i> -neohexaose
TFLNH(a)	Trifucosyl-lacto- <i>N</i> -hexaose (a)
TriLacNAc	Tri- <i>N</i> -acetyl- <i>D</i> -lactosamine
PA	Pyridylamino- or 2-aminopyridine
MS	Mass spectrometry
TEA	Trimethylamine
FA	Formic acid
QC	Quality control
LOQ	Limit of quantification
EIC	Extracted ion chromatogram

Tumor markers, along with diagnostic imaging, are important tools for the clinical diagnosis of cancer, as well as for the evaluation of treatment efficacy and monitoring for possible recurrence. Several glycan tumor markers, including CA19-9, sialyl Tn, sialyl Lewis X and Du-Pan-2, have been developed and are widely used in clinical practice^{1–5}. However, because many cases are often negative for any marker despite disease progression, further novel marker development will undoubtedly improve diagnosis.

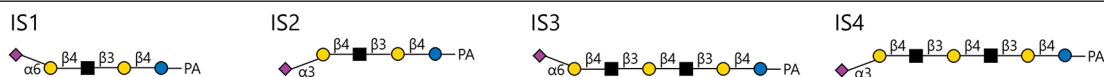
Various approaches, including mass spectrometry-based or lectin-based methods, have been employed in the search for novel glycan tumor markers^{6–9}. We have been conducting detailed glycan structural analysis of human cancer tissues to identify novel glycan tumor markers. These studies revealed an abnormal accumulation of free-glycans in cancer tissues^{10,11}. Specifically, a large amount of NeuAc-linked free complex-type *N*-glycans were found to accumulate in pancreatic cancers¹⁰. Furthermore, free-*N*-glycans containing an unusual sialic acid, 2-keto-3-deoxy-nonulosonic acid, as well as common sialic acid, NeuAc, were found in prostate cancer specimens¹¹. These results suggest that free-glycans are potential tumor markers. Because free-glycans have a relatively small molecular weight, they are rapidly excreted in urine. Urinary biomarkers offer a non-invasive and readily accessible means of monitoring the health of a patient. Therefore, free-glycan marker discovery studies have been performed using patient urine samples^{12–15}.

The urinary free-glycans were divided into the following three fractions and their glycan structures studied in detail: 1) small acidic glycans, mainly composed of lactose-core glycans, LacNAc-core glycans, and mono-antennary free-*N*-glycans, (abbreviated as small acidic)¹², 2) large acidic glycans, mainly composed of multi-antennary *N*-glycans (abbreviated as large acidic)¹³, and 3) neutral glycans^{14,15}. In all, 135, 81, and 78 free-glycans were identified in the small acidic, large acidic, and neutral fractions, respectively. Among them, many marker candidates were found to be elevated in cancer patients that can be characterized into eight groups according to their backbone structure/modification as follows: 1) sialyl lactose-core glycans, 2) glycans having type1 structure, 3) glycans with an unusual structure, newly identified in our previous study, including Fucα1-3Xylα1-3Glc glycans and *D*-arabinose containing glycan, 4) sulfated glycans, 5) mono-, 6) di-, 7) tri-antennary Gn1-core free-*N*-glycans, and 8) multi-antennary Gn2-core free-*N*-glycans. Here, the clinical usefulness of the marker candidates was investigated. In particular, the following three issues were addressed.

- 1) Improvement of accuracy: Candidate markers were identified by HPLC-based methods, and their quantification was initially performed by selected reaction monitoring (SRM). However, in the absence of internal standards (ISs) the quantification was only semi-quantitative, which is problematic for clinical applications. It is impractical to synthesize ISs, especially stable-isotope labeled ISs, for each of the many free-glycan markers. Instead, we synthesized four unique lactose-core glycans containing GlcNGc (*N*-glycolylglucosamine), which are not present in human (listed in Table 1). We then examined whether these glycans can act as ISs for various kinds of free-glycans. Assays were evaluated using eight commercially available reference glycans, which are present in human (Supplementary Table 1), and 4ISs.
- 2) Selection of useful markers: Among the markers that can be quantified with high accuracy in urine samples from most cases, a total of 12 markers (numbered and listed in Table 1) were chosen. These include three sialyl lactose-core glycans, three sulfated glycans, four Gn1-core free-*N*-glycans, and two glycans with an unusual structure. The chosen markers were subjected to SRM analysis. Univariate and multivariate receiver operating characteristic (ROC) curve analyses were performed to investigate the effectiveness of the individual marker alone or a combination of multiple markers for cancer diagnosis.
- 3) Increase the number of cases analyzed and diseases targeted: Previous exploratory studies were conducted using urine obtained from a small number of patients with gastric, colorectal, and pancreatic cancer, as well as healthy subjects. To improve statistical analysis, the number of cases analyzed needs to be increased. Furthermore, there is no information on the efficacy of these markers for other carcinomas. In this study, in addition to gastrointestinal cancers, various malignancies including lung cancer, esophagus cancer, malignant lymphoma, and sarcoma were added to the list of diseases to be analyzed (Table 2).

In summary, 12 free-glycan tumor marker candidates were simultaneously measured in urine samples from 623 cases, 120 healthy subjects and 503 patients with malignant tumors. Samples were analyzed by SRM using ISs, and their clinical usefulness investigated. Using a combination of four markers, area under curve (AUC) values of 0.910, 0.867, and 0.914 were obtained for gastric, colorectal, and pancreatic cancer, respectively. Furthermore, these markers were elevated in various malignancies including sarcoma, suggesting that they could also be used to monitor a range of malignancies.

Internal Standards



Free-glycan markers

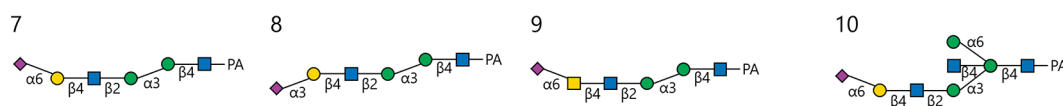
Lactose-core glycans



Sulfated glycans



Gn1-core free-*N*-glycans



Unusual glycans



Table 1. Structures of the internal standards and the free-glycan markers.

The abbreviations and number of each free-glycan marker are shown above the structure. Monosaccharide symbols are indicated as follows: blue circle, Glc; blue square, GlcNAc; green circle, Man; yellow circle, Gal; red triangle, Fuc; purple diamond, sialic acid (NeuAc); orange star, xylose; green star, arabinose; S, sulfate; black square, GlcNGc.

[illegible]

Table 2. Clinical information on the patients. N, normal control; G, gastric cancer; C, colorectal cancer; P, pancreatic cancer; B, bile duct cancer; E, esophagus cancer; L, lung cancer; H, hematological tumor; Ov, ovarian cancer; En, endometrial cancer; Ce, cervical cancer; Sa, sarcoma. M and F refer to male and female, respectively.

Methods

Urine samples

Urine samples of patients with gastric cancer (n=140), colorectal cancer (n=87), pancreatic cancer (n=55), bile duct tumor (n=13), esophageal cancer (n=45), lung cancer (n=29), various types of hematologic tumor (n=31), ovarian cancer (n=12), endometrial cancer (n=13), cervical cancer (n=9), and various types of sarcoma (n=69) were obtained from Osaka International Cancer Institute (Table 2). Additional information on the number of cases in each stage of disease progression alongside the histological classification is shown in Supplementary Table 2. Urine samples of normal controls (n=120) were obtained from cancer-free healthy volunteers. This study was approved by the Local Ethics Committee of Osaka International Cancer Institute.

Informed consent was obtained from each patient and volunteer. All experiments were conducted in accordance with relevant guidelines and regulations.

Internal standard glycans and reference glycans

The four different types of free-glycans used as ISs were synthesized by Tokyo Kasei Industry (Tokyo, Japan) (Table 1). These glycans were α 2–6 or α 2–3 sialylated lacto-*N*-neotetraose, and α 2–6 or α 2–3 sialylated straight type of lacto-*N*-neohexaose, with the GlcNAc (*N*-acetylglucosamine) residue replaced by GlcNGc (*N*-glycolylglucosamine), i.e. NeuAca2-6Gal β 1-4GlcNGc β 1-3Gal β 1-4Glc, NeuAca2-3Gal β 1-4GlcNGc β 1-3Gal β 1-4Glc, NeuAca2-6Gal β 1-4GlcNGc β 1-3Gal β 1-4GlcNGc β 1-3Gal β 1-4Glc, and NeuAca2-3Gal β 1-4GlcNGc β 1-3Gal β 1-4GlcNGc β 1-3Gal β 1-4Glc, and designated as IS1, IS2, IS3, and IS4, respectively. One GlcNAc unit was replaced by GlcNGc in IS1 and IS2, and two GlcNAc units were replaced by GlcNGc in IS3 and IS4. The glycans are depicted according to the “Symbol Nomenclature for Glycans”¹⁶. GlcNGc that is not registered in the Symbol Nomenclature is marked as a black square. Stock solutions of the four ISs were prepared by mixing them so that each had a concentration of 10 pmol/L.

Eight reference glycans were purchased to use in the urinary free-glycan assay validation (Supplementary Table 1). Sialyl Lewis A tetra (SLA) was purchased from Dextra Laboratories (Reading, UK), LST-b and LST-c were from Agilent (Santa Clara, CA, USA), lacto-*N*-neohexaose (LNnH) was from Biosynth (Compton, UK), difucosyl-lacto-*N*-hexaose (a) (DFLNH(a)) and difucosyl-lacto-*N*-neohexaose (DFLNnH) were from IsoSep AB (Tullinge, Sweden), trifucosyl-lacto-*N*-hexaose (a) (TFLNH(a)) was from Sigma-Aldrich (St Louis, MO, USA), tri-*N*-acetyl-D-lactosamine (TriLacNAc) was from Elicityl (Crolles, France). Stock solutions of the eight reference glycans were prepared by mixing them so that each had a concentration of 100 pmol/L.

Isolation and labeling of free-glycans from urine samples

To normalize the concentration of urine, volumes used for sample preparation were corrected by creatinine concentration. Urine samples equivalent to 200 μ g creatinine were spiked with 10 pmol ISs and used as starting material. The urine samples were applied to a column packed with 1 mL of Dowex 50W-X8 resin (H^+ -form, 100–200 mesh; FUJIFILM Wako Pure Chemical, Osaka, Japan) followed by 4 mL water. Flowthrough and wash fractions were collected. Eluates were neutralized by addition of 0.4 mL of saturated sodium bicarbonate prior to loading onto a graphite carbon cartridge (InertSepGC 300 mg; GL Science, Tokyo, Japan). The column was washed with 5 mL water and bound oligosaccharides were then eluted with 5 mL of 60% acetonitrile in 0.1% trifluoroacetic acid, and lyophilized¹². Labeling of free-glycan with aminopyridine (PA-glycan) was performed as previously reported^{12,17}. Oligosaccharides were incubated with 40 μ L of a coupling reagent (prepared by mixing 552 mg 2-aminopyridine and 200 μ L acetic acid) at 90 °C for 60 min. The resulting Schiff base was reduced at 80 °C for 35 min in the presence of 140 μ L of a freshly prepared mixture of 200 mg dimethylamine–borane complex, 50 μ L water, and 80 μ L acetic acid. The reaction mixture was diluted with 750 μ L water and extracted twice with 800 μ L water-saturated phenol:chloroform (1:1, v/v) and once with 800 μ L chloroform. The aqueous phase was applied to a column packed with 1 mL of Dowex 50W-X8 (NH_4^+ -form, 50–100 mesh, FUJIFILM Wako Pure Chemical), which was washed with 3.2 mL of 20 mM ammonium acetate buffer, pH 8.5. Eluates were then extracted using a graphite carbon cartridge (InertSepGC 150 mg; GL Science). The column was washed with 3 mL of 50 mM ammonium acetate, pH 7.0, and bound PA-glycans were eluted with 3 mL of 60% acetonitrile in 50 mM ammonium acetate, pH 7.0. PA-glycans were lyophilized and dissolved in 400 μ L water.

Mass spectrometry for quantification by SRM

Quantification by SRM was performed on an LC–MS system comprising a Shimadzu LC-20A HPLC system (Shimadzu, Kyoto, Japan) linked to 4500 Q Trap hybrid triple quadrupole/linear ion trap mass spectrometer with a Turbo V ion source (AB SCIEX, Framingham, MA, USA). Twelve markers were measured in triethylamine-based (TEA-based) solvents. For assay validation, conditions commonly used for LC/MS employing formic acid-based (FA-based) solvents were also considered. Reversed-phase HPLC conditions for TEA-based and FA-based solvents are listed in Supplementary Table 3. The quality of the assay was assessed using a mixture of analysis samples from healthy subjects and cancer patients that were prepared and measured in intervals of 10 samples^{18,19}. The same mass spectrometer settings were used under both HPLC conditions with TEA-based and FA-based solvents: positive-ion mode, an ion spray voltage of 5 kV, curtain gas of 30 psi, nebulizer gas (GS1) of 65 psi, turbo gas (GS2) of 55 psi and an interface heater temperature of 500 °C. The settings for SRM, including Q1 and Q3 values and collision energy are shown in Supplementary Table 4. The data were recorded by Analyst 1.7.2 software and analyzed by SCIEX OS 2.0.0 software (AB SCIEX).

Validation procedures

Eight reference glycans and four ISs were used to validate this assay under both TEA-based and FA-based HPLC conditions. Calibrators and quality control (QC) samples were prepared from pooled control urine, which were subdivided into aliquots equivalent to 200 μ g creatinine. These samples were spiked with either 300 pmol, 100 pmol, 30 pmol, 10 pmol, 3 pmol, 1 pmol, or 0.3 pmol of the eight reference glycans. Likewise, QCs were prepared by spiking control urine samples with either 20 pmol (High), 5 pmol (Medium), or 1 pmol (Low) of the eight reference glycans. Calibration samples and QCs were treated in the same manner as study samples after the addition of 10 pmol IS mixture. The ratio of the area of reference glycans to that of the IS was converted to concentration by means of a calibration curve. The contribution of the endogenous reference glycans in control urine used for calibration was eliminated by subtracting the peak area ratio of each analyte to internal standard of the blank from all of the other points.

Statistical analysis

Univariate and multivariate analyses were performed on the web-based platform MetaboAnalyst version 6.0²⁰ and JMP statistical software Version 17.

Results

Validation of method

Initially, we investigated whether the four synthesized glycans can function as ISs for various free-glycans present in the samples. Validation was performed using eight reference molecules that are present as free-glycans in urine ranging from minor to major components. Reference molecules varied in length from four to nine glycans with different backbones/modifications. Control urine was spiked with 10 pmol of IS mixture and 0.3–300 pmol of reference glycan mixture. Free-glycans were then isolated, subjected to pyridylation, and analyzed by LC-SRM under LC conditions for both TEA- and FA-based solvents. In all, 64 calibration curves were generated with four different ISs versus eight reference glycans under two sets of LC conditions. The linearity of each calibration curve was within acceptable limits (i.e., mean coefficient of determination (r^2) > 0.99 ($n = 5$)). Details of the statistical analysis are given in Supplementary Table 5. As in the linearity experiment, we determined the accuracy and precision of the method at low (1 pmol), medium (5 pmol), and high (20 pmol) concentrations of the eight reference glycans ($n = 5$ for each concentration). The results are summarized in Supplementary Table 6. With minor exceptions, the accuracy and mean coefficient of variation for the eight reference glycans measured using four different ISs under two different HPLC conditions were < 15% for the intraday ($n = 5$) and interday ($n = 5$) samples. These results demonstrated that the eight reference glycans with diverse structures could be accurately quantified using any of the four ISs. Therefore, the four ISs were judged to be appropriate for the quantification of free-glycans. Of the four, IS1 and IS3 were used for the actual sample measurements. Unlike the reference glycans, none of the 12 markers are commercially available. Due to difficulties in obtaining sufficient quantities of these markers it was not possible to generate a standard curve. Consequently, absolute quantities could not be measured. Quantification of these markers was based on the relative signal area ratio of marker to IS (Marker/IS).

Quantification of glycan markers by reversed-phase HPLC-SRM

Ionization efficiency differs depending on the structure of the glycan. Thus, the exact limit of quantification (LOQ) of the 12 markers could not be determined. Nonetheless, calibration curves of the eight reference glycans displayed high linearity in the range 0.3–300 pmol. Instead of intact marker glycans, four distinct PA-labeled markers (#3, #4, #5, #7) were obtained from previous studies. PA-labeled free-glycan mixtures, including 10 pmol IS, were spiked with 0.1, 0.3, 1, 3, or 10 pmol of the PA-labeled markers and subjected to SRM measurement. High linearity ranging from 0.1–10 pmol was observed. Thus, area ratio of Marker/IS of 0.01 was tentatively set as LOQ for this study.

Markers that in many subjects showed values below the LOQ or that could not be accurately quantified due to insufficient separation from isomer were excluded from this analysis. Twelve free-glycan markers were selected for quantification of a large number of samples to assess their clinical utility (Table 1). The markers were classified into four categories according to their structural characteristics, i.e. lactose-core glycans (#1, #2, #3), sulfated glycans (#4, #5, #6), mono-antennary Gn1-core glycans (#7, #8, #9, #10), and unusual glycans (#11, #12) (Table 1). The markers in urine samples from healthy subjects or patients with various malignant tumors were measured simultaneously by TEA-based LC-SRM, (Fig. 1 and Fig. 2). Representative overlays of extracted ion chromatograms (EIC) of two ISs (IS1 and IS3), seven high abundant markers (#1, #2, #4, #5, #7, #8, #9), and five relatively low abundant markers (#3, #6, #10, #11, #12) from normal cases are shown in Fig. 1A, 1B, and 1C, respectively. Because the intensity of IS1 is about twice that of IS3, IS1 was used for quantification (Fig. 1A). All peaks were either a single major peak or clearly separated from the corresponding isomer. Even relatively low abundant markers were detected as distinct peaks (signal-to-noise ratios > 10).

The levels of #2, #4, #5, #7, #8, and #11 for healthy subjects and patients with various malignant tumors are shown in Fig. 2, and those of #1, #3, #6, #9, #10, and #12 are shown in Supplementary Fig. 1. Six (#5, #7, #9, #10, #11, #12) of the 12 glycans were present in samples from cancer patients at significantly elevated levels compared to normal controls. These glycans were expressed on a logarithmic 10 scale (#5, #7 and #11 in Fig. 2; #9, #10, #12 in Supplementary Fig. 1). Glycans classified into four groups show similar profiles, so quantitative results are described for each group.

Sialyl lactose-core glycans: #1 is a simple α 2–3 sialylated lactose-core glycan. #2 and #3 have sialyl Lewis X structure. The levels of #1, #2, and #3 were moderately elevated in most cancers including gastrointestinal, pancreatic, bile duct, and lung cancers as well as sarcoma. Prominent elevation of #1 and #3 were observed in sarcoma patients.

Sulfated glycans: #4, #5 and #6 are 6-sulfo-lactose-core glycans. #6 also possesses type1 structure, sialyl Lewis A extension. All the sulfated glycans showed a marked increase in pancreatic and bile duct cancer patients. In addition to pancreatic and bile duct cancer patients #4 and #5 were moderately elevated in esophageal cancer patients.

Mono-antennary Gn1-core free-N-glycans: #7, #8, #9, and #10 are Man₂GlcNAc₁-core glycans containing α 2,6-sialyl LacNAc, α 2,3-sialyl LacNAc, α 2,6-sialyl LacdiNAc, and bisecting GlcNAc, respectively. These four glycans tends to be prominently elevated in gastric cancer patients.

Unusual glycans: #11 is a terminal sialylated and fucosyl-xylose containing glycan, NeuAc α 2-3Gal β 1–4(Fuc α 1–3)Xyl α 1–3Glc. #12 is a D-arabinose-containing free-N-glycan. Elevated levels of these unusual glycans was observed in most malignant tumors including sarcoma.

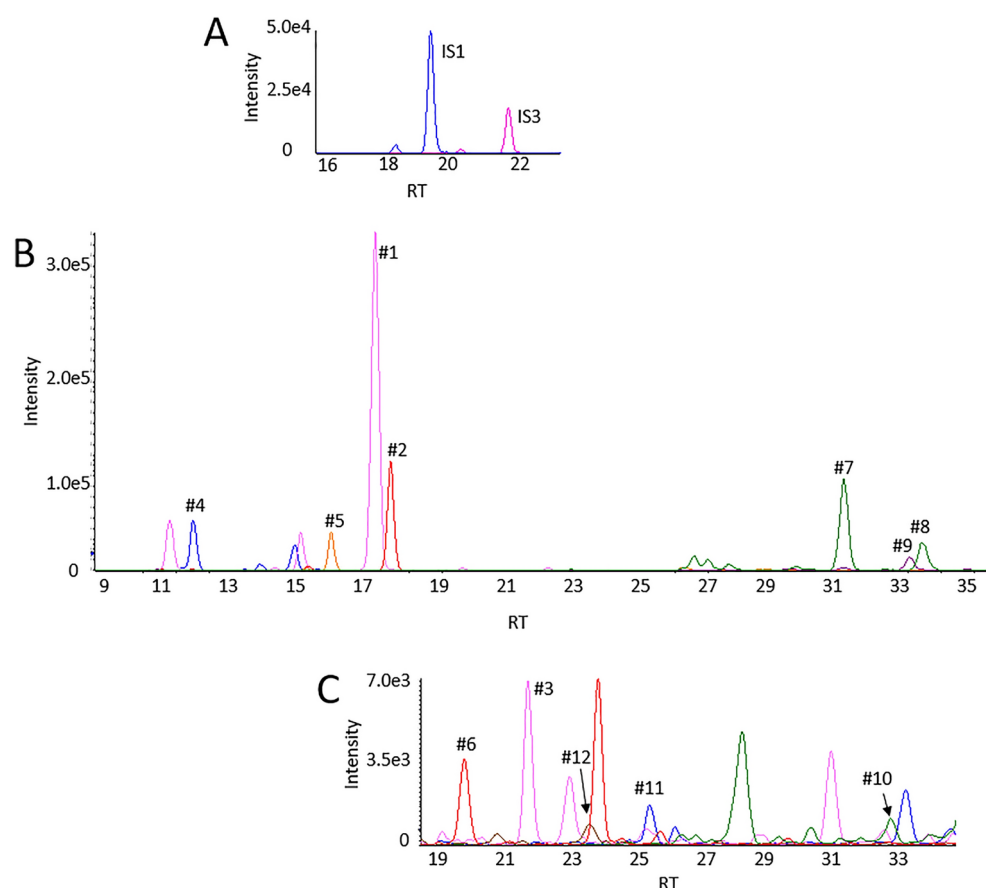


Fig. 1. Overlaid extracted ion chromatogram of internal standards and marker glycans. The representative EIC from the normal subject is shown. (A), IS1 (blue) and IS3 (red); (B), high abundant glycans, #1 (pink), #2 (red), #4 (blue), #5 (orange), #7 (green), #8 (green), #9 (brown); (C), relatively low abundant glycans, #3 (pink), #6 (red), #10 (green), #11 (blue), #12 (brown).

Univariate ROC curve analyses

Univariate ROC curve analyses were performed to assess the usefulness of each marker alone for cancer diagnosis (Table 3). AUC and 95% confidence interval (CI) for each marker in gastric, colorectal, and pancreatic cancers are shown. The highest AUCs for gastric, colorectal, and pancreatic cancers were 0.732 for #12, 0.805 for #11, and 0.854 for #11, respectively. These results indicate that individual markers may not be sufficient to diagnose cancer.

Exploratory multivariate ROC analysis

A multivariate ROC curve analysis using a linear support vector machine algorithm was performed to determine the appropriate number of features to improve cancer diagnostic performance (Supplementary Fig. 2). Sufficient AUC was thought to be obtained for gastric, colorectal, and pancreatic cancer in the four features model. Four significant features were selected based on the forward stepwise selection method (see Fig. 3). The four selected markers are detailed below. Gastric cancer; two types of sulfated glycans #4 and #5, two mono-antennary Gn1-core free-N-glycans, #7 and #8. The four markers selected were identical for colorectal and pancreatic cancers; lactose-core glycan #2, two mono-antennary Gn1-core free-N-glycans, #7 and #8, and unusual glycan #11. To assess the utility of using a combination of the four markers, ROC curve analysis with the logistic regression algorithm and ten-fold cross-validation was performed. The AUC of the combination of the four markers was 0.910, 0.867, and 0.914 for gastric, colorectal, and pancreatic cancer, respectively. In gastric cancer, the highest AUC among the four markers was 0.695 for #7, but a combination of markers considerably increased the AUC to 0.910, reflecting a significant improvement. For colorectal and pancreatic cancer, #11 alone showed AUC values of 0.805 and 0.854, respectively. However, a combination of the four markers for colorectal and pancreatic cancer showed a slight increase in performance to 0.867 and 0.914, respectively. The likely reason for the seemingly low diagnostic improvement with the combination of markers is because even #11 alone has a high AUC (>0.8). The results of the ROC curve analysis of gastric and colorectal cancer by stage (stage I + II, stage III, and stage IV) are presented in Supplementary Table 7. AUC values increased with disease stage.

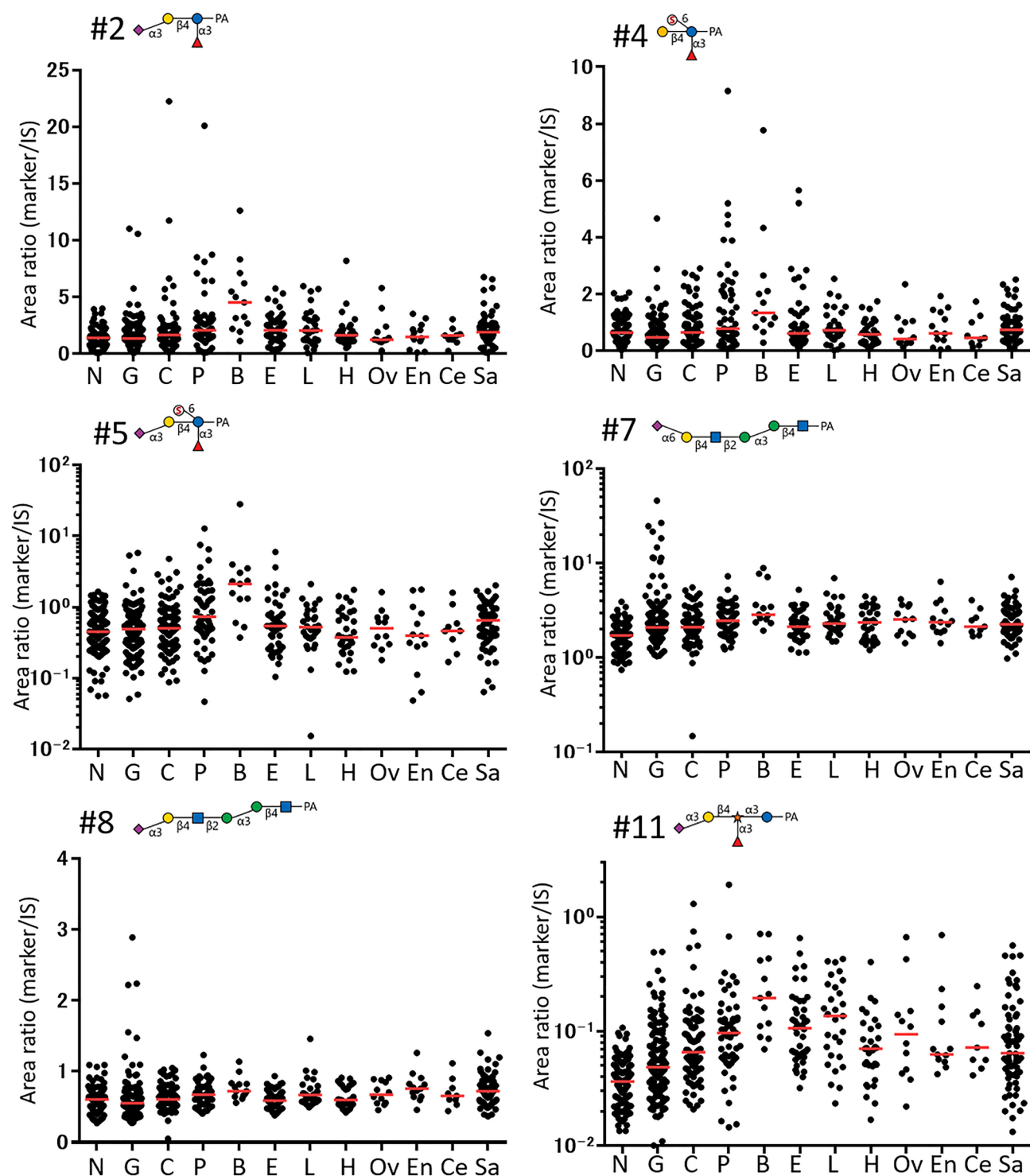


Fig. 2. Levels of free-glycan markers in the urine from SRM analysis. Levels of six glycans (#2, #4, #5, #7, #8, #11) are shown. Abbreviations of disease names are the same as in Table 2. The median is indicated by the red line. The levels of #5, #7 and #11 are expressed on a logarithmic 10 scale.

Discussion

Among the large number of urinary free-glycan markers identified in our previous studies, 12 markers were selected, and their clinical utility investigated. These detailed analyses revealed that a combination of markers improved the diagnostic performance for gastric, colorectal, and pancreatic cancers. The levels of these markers were also elevated in patients with a range of malignant tumors, indicating their broad clinical utility.

To precisely quantify urinary free-glycan markers, we have established a method for the quantification of free-glycans in urine by SRM using ISs containing GlcNGc. The presence of GlcNGc containing glycans has not

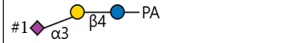
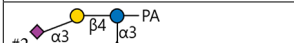
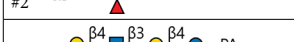
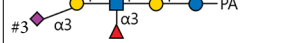

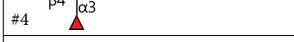
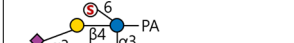
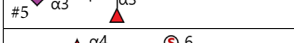

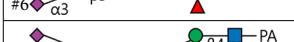
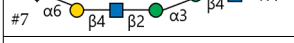
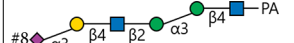
	Gastric	Colorectal	Pancreas
#1 	0.704 (0.641 ~ 0.766)	0.669 (0.594 ~ 0.743)	0.762 (0.682 ~ 0.843)
#2 	0.521 (0.450 ~ 0.591)	0.607 (0.527 ~ 0.687)	0.687 (0.589 ~ 0.784)
#3 	0.640 (0.572 ~ 0.707)	0.714 (0.644 ~ 0.785)	0.752 (0.671 ~ 0.833)
#4 	0.588 (0.518 ~ 0.657)	0.498 (0.417 ~ 0.580)	0.607 (0.504 ~ 0.709)
#5 	0.525 (0.454 ~ 0.596)	0.523 (0.441 ~ 0.605)	0.683 (0.590 ~ 0.776)
#6 	0.568 (0.498 ~ 0.638)	0.619 (0.540 ~ 0.698)	0.585 (0.482 ~ 0.688)
#7 	0.695 (0.632 ~ 0.758)	0.690 (0.616 ~ 0.763)	0.795 (0.721 ~ 0.869)
#8 	0.435 (0.365 ~ 0.505)	0.527 (0.448 ~ 0.607)	0.640 (0.553 ~ 0.727)
#9 	0.653 (0.587 ~ 0.719)	0.600 (0.521 ~ 0.678)	0.709 (0.625 ~ 0.793)
#10 	0.706 (0.643 ~ 0.768)	0.620 (0.541 ~ 0.699)	0.780 (0.699 ~ 0.862)
#11 	0.665 (0.600 ~ 0.730)	0.805 (0.744 ~ 0.866)	0.854 (0.780 ~ 0.928)
#12 	0.732 (0.671 ~ 0.792)	0.767 (0.698 ~ 0.835)	0.776 (0.694 ~ 0.858)

Table 3. Univariate receiver operating characteristic (ROC) curve analysis. Area under curve (AUC) and 95% confidence interval (CI) for each marker in gastric, colorectal, and pancreatic cancers are shown.

been reported. These GlcNGc-containing glycans functioned well as ISs in our assay procedure with less than detection sensitivity in urine.

The markers found at elevated levels in cancer patients were characterized into eight groups according to their structure and modification: 1) sialylated lactose-core glycans, 2) glycans having type1 structure, 3) unusual glycans, 4) sulfated glycans, 5) mono-, 6) di-, 7) tri-antennary Gn1-core free-*N*-glycans, and 8) multi-antennary Gn2-core free-*N*-glycans. In the present analysis, we focused on 12 quantifiable markers belonging to the following four groups; sialylated lactose-core, sulfated glycans, unusual glycan, mono-antennary Gn1-core free-*N*-glycans. In addition to these 12 markers, the levels of 32 markers, selected from all eight groups outlined above, were measured. However, these markers were excluded from the detailed statistical analysis because many cases were either below the LOQ, or there was insufficient separation from isomer. Nonetheless, most of these glycans were clearly elevated in cancer patients and showed distinct profiles from the 12 markers. It is important to improve detection sensitivity and chromatographic resolution so that excluded markers can be accurately quantified. As more markers are added to the combination, the accuracy of the diagnosis may be further improved. Moreover, novel markers have been identified in the subsequent search. Additional analyses of these new markers will also be necessary.

When we examined the utility of the 12 markers, an individual marker alone was found to be unsuitable for accurate diagnosis. A combination of four markers was required to diagnose gastric, colorectal, and pancreatic cancers. Intriguingly, the four markers selected to diagnose colorectal and pancreatic cancers were identical. Moreover, #7 and #8 were common to the diagnosis of gastric, colorectal, and pancreatic cancer. Thus, of the 12 markers, only six were required to diagnose the three forms of cancer. These six markers include members from all four groups, suggesting that a combination of glycans with various structural characteristics is preferable. In addition to the structural diversity of the selected markers, the significant improvement in diagnosis can be attributed to the different mechanisms that underlie their elevation in cancer. Specifically, four of the selected glycans are biosynthesized via pathways that are enhanced in cancer tissues (#4 and #5 of gastric cancer, #2 and #11 of colorectal and pancreatic cancer). The other two glycans are derived as a result of lysosomal dysfunction (#7, #8 of three cancers). #7 and #8, which were selected as markers for three types of cancer, are major degradation products of *N*-glycans. It has been reported that these glycans markedly accumulate in gastric cancer-derived cell lines due to the low integrity of lysosomal membranes²¹. Moreover, our glycan structural analysis of human gastric, colorectal, pancreatic, and prostate cancers showed #7 and #8 to be the most abundant free-glycans (data from gastric and colorectal cancers are unpublished)^{10,11}. Consequently, #7 and #8 may be useful markers for other types of cancer that develop lysosomal dysfunction as they progress. Unlike #7 and #8, the biosynthesis of glycans #2, #4, #5 and #11 are elevated in cancer cells. During malignant transformation the levels of both

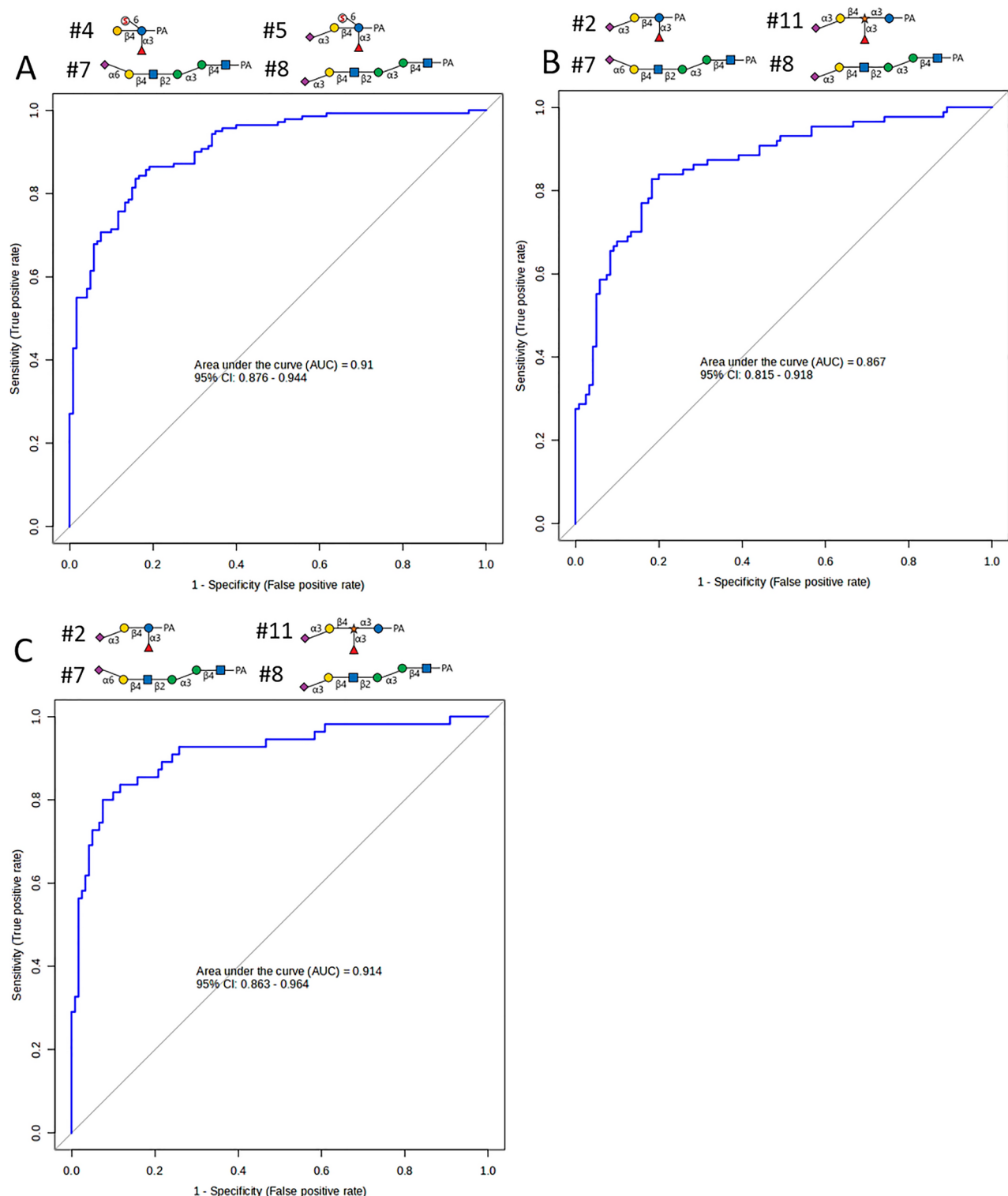


Fig. 3. Multivariate receiver operating characteristic (ROC) curve analysis. ROC curve evaluation of four marker models of gastric (A), colorectal (B) and pancreatic (C) cancer patients. Four markers selected by a stepwise selection method are indicated above the ROC curve.

sialylation and fucosylation are enhanced^{22,23}. Indeed, the activities of sialyltransferases, and 6-sulfotransferases in human cancer cells were markedly increased compared to those of normal mucosal epithelial cells²⁴ (data of 6-sulfotransferases is unpublished). #2 and #11 undergo both α 2/3-sialylation and α 1/3-fucosylation, forming sialyl Lewis X structures that are well known tumor-associated carbohydrate antigens. Similarly, #4 and #5 undergo both 6-sulfation and α 1/3-fucosylation to generate 6-sulfo Lewis X structures. Glycans with 6-sulfo

Lewis X structure have been found to accumulate in human colon cancers²⁵. The combination of markers resulted in a particularly significant increase in AUC for gastric cancer, which may be attributed to #7 and #8 being specifically elevated in gastric cancers (Fig. 2).

The expansion of omics research has driven many attempts to use a combination of multiple markers for the diagnosis and prognosis of various diseases, including cancer^{26–31}. While proteins, circulating DNA, and lipids have been the main targets of research to date, we believe the focus of attention should also include glycans.

To investigate the potential of free-glycans as markers in other malignancies, 12 markers were also quantified in a small number of cases with various malignancies. None of the glycans were elevated in specific malignancies, such as those observed in gastric and pancreatic cancer. Instead, glycans elevated in almost all malignancies, including various types of malignant lymphoma and sarcoma, were observed (e.g., #10, #11, #12). In addition to tumor markers that are specific for particular forms of cancer, we believe urinary free-glycans could also serve as valuable markers for a variety of cancers.

Data availability

All data generated or analyzed during the study are included either in this published article and its supplementary information file or can be made available from the corresponding author on reasonable request.

Received: 7 November 2024; Accepted: 13 March 2025

Published online: 24 March 2025

References

1. Thomas, D., Rathinavel, A. K. & Radhakrishnan, P. Altered glycosylation in cancer: A promising target for biomarkers and therapeutics. *Biochim. Biophys. Acta Rev. Cancer* **1875**, 188464. <https://doi.org/10.1016/j.bbcan.2020.188464> (2021).
2. Pinho, S. S. & Reis, C. A. Glycosylation in cancer: Mechanisms and clinical implications. *Nat. Rev. Cancer* **15**, 540–555. <https://doi.org/10.1038/nrc3982> (2015).
3. Chia, D. et al. Use of monoclonal antibodies to sialylated Lewisx and sialylated Lewisy for serological tests of cancer. *Cancer Res.* **45**, 435–437 (1985).
4. Takasaki, H. et al. Correlative study on expression of CA 19–9 and DU-PAN-2 in tumor tissue and in serum of pancreatic cancer patients. *Cancer Res.* **48**, 1435–1438 (1988).
5. Del Villano, B. C. et al. Radioimmunoassay for a monoclonal antibody-defined tumor marker, CA 19–9. *Clin. Chem.* **29**, 549–552 (1983).
6. McDowell, C. T. et al. Imaging mass spectrometry and lectin analysis of N-Linked Glycans in Carbohydrate Antigen-Defined Pancreatic Cancer Tissues. *Mol. Cell. Proteomics* **20**, 100012. <https://doi.org/10.1074/mcp.RA120.002256> (2021).
7. Dotz, V. & Wuhler, M. N-glycome signatures in human plasma: Associations with physiology and major diseases. *FEBS Lett.* **593**, 2966–2976. <https://doi.org/10.1002/1873-3468.13598> (2019).
8. Marrero Roche, D. E. & Chandler, K. B. Clinical glycoprotein mass spectrometry: The future of disease detection and monitoring. *J. Mass Spectrom.* **59**, e5083. <https://doi.org/10.1002/jms.5083> (2024).
9. Peng, W. et al. MS-based glycomics and glycoproteomics methods enabling isomeric characterization. *Mass Spectrom. Rev.* **42**, 577–616. <https://doi.org/10.1002/mas.21713> (2023).
10. Yabu, M. et al. Accumulation of free Neu5Ac-containing complex-type N-glycans in human pancreatic cancers. *Glycoconj. J.* **30**, 247–256. <https://doi.org/10.1007/s10719-012-9435-9> (2013).
11. Yabu, M. et al. Occurrence of free deaminoneuraminic acid (KDN)-containing complex-type N-glycans in human prostate cancers. *Glycobiology* **23**, 634–642. <https://doi.org/10.1093/glycob/cws132> (2013).
12. Hanzawa, K. et al. Investigation of acidic free-glycans in urine and their alteration in cancer. *Glycobiology* **31**, 391–409. <https://doi.org/10.1093/glycob/cwaa100> (2021).
13. Hanzawa, K. et al. Increased levels of acidic free-N-glycans, including multi-antennary and fucosylated structures, in the urine of cancer patients. *PLoS One* **17**, e0266927. <https://doi.org/10.1371/journal.pone.0266927> (2022).
14. Tanaka-Okamoto, M., Hanzawa, K., Murakami, H., Mukai, M. & Miyamoto, Y. Identification of beta1-3 galactosylglucose-core free-glycans in human urine. *Anal. Biochem.* **641**, 114427. <https://doi.org/10.1016/j.ab.2021.114427> (2022).
15. Tanaka-Okamoto, M. et al. Occurrence of a D-arabinose-containing complex-type free-N-glycan in the urine of cancer patients. *Sci. Rep.* **12**, 4889. <https://doi.org/10.1038/s41598-022-08790-0> (2022).
16. Varki, A. et al. Symbol nomenclature for graphical representations of glycans. *Glycobiology* **25**, 1323–1324. <https://doi.org/10.1093/glycob/cwv091> (2015).
17. Natsuka, S., Hirohata, Y., Nakakita, S., Sumiyoshi, W. & Hase, S. Structural analysis of N-glycans of the planarian *Dugesia japonica*. *FEBS J.* **278**, 452–460. <https://doi.org/10.1111/j.1742-4658.2010.07966.x> (2011).
18. Want, E. J. et al. Global metabolic profiling of animal and human tissues via UPLC-MS. *Nat. Protoc.* **8**, 17–32. <https://doi.org/10.1038/nprot.2012.135> (2013).
19. Want, E. J. et al. Global metabolic profiling procedures for urine using UPLC-MS. *Nat. Protoc.* **5**, 1005–1018. <https://doi.org/10.1038/nprot.2010.50> (2010).
20. Ewald, J. D. et al. Web-based multi-omics integration using the Analyst software suite. *Nat. Protoc.* **19**, 1467–1497. <https://doi.org/10.1038/s41596-023-00950-4> (2024).
21. Ishizuka, A. et al. Accumulation of free complex-type N-glycans in MKN7 and MKN45 stomach cancer cells. *Biochem. J.* **413**, 227–237. <https://doi.org/10.1042/BJ20071562> (2008).
22. Hakomori, S. Glycosylation defining cancer malignancy: New wine in an old bottle. *Proc. Natl. Acad. Sci. U.S.A.* **99**, 10231–10233 (2002).
23. Fukushi, Y., Nudelman, E., Levery, S. B., Hakomori, S. & Rauvala, H. Novel fucolipids accumulating in human adenocarcinoma. III. A hybridoma antibody (FH6) defining a human cancer-associated difucoganglioside (VI3NeuAcV3III3Fuc2nLc6). *J. Biol. Chem.* **259**, 10511–10517 (1984).
24. Misonou, Y. et al. Comprehensive clinico-glycomic study of 16 colorectal cancer specimens: elucidation of aberrant glycosylation and its mechanistic causes in colorectal cancer cells. *J. Proteome Res.* **8**, 2990–3005. <https://doi.org/10.1021/pr900092r> (2009).
25. Shida, K. et al. Unusual accumulation of sulfated glycosphingolipids in colon cancer cells. *Glycobiology* **19**, 1018–1033. <https://doi.org/10.1093/glycob/cwp083> (2009).
26. Kim, H. et al. Development of a Fit-For-Purpose Multi-Marker Panel for Early Diagnosis of Pancreatic Ductal Adenocarcinoma. *Mol. Cell. Proteomics* **23**, 100824. <https://doi.org/10.1016/j.mcpro.2024.100824> (2024).
27. Zhang, X. et al. Integrative multianalytical model based on novel plasma protein biomarkers for distinguishing lung adenocarcinoma and benign pulmonary nodules. *J. Proteome Res.* **23**, 277–288. <https://doi.org/10.1021/acs.jproteome.3c00551> (2024).

28. Zieren, R. C. et al. Diagnostic liquid biopsy biomarkers in renal cell cancer. *Nat. Rev. Urol.* **21**, 133–157. <https://doi.org/10.1038/s41585-023-00818-y> (2024).
29. Xie, Y. et al. DNA Methylation-Based Testing in Peripheral Blood Mononuclear Cells Enables Accurate and Early Detection of Colorectal Cancer. *Cancer Res.* **83**, 3636–3649. <https://doi.org/10.1158/0008-5472.CAN-22-3402> (2023).
30. Krishnan, S. T. et al. Staging of colorectal cancer using lipid biomarkers and machine learning. *Metabolomics* **19**, 84. <https://doi.org/10.1007/s11306-023-02049-z> (2023).
31. Katoh, K. et al. Serum Free Fatty Acid Changes Caused by High Expression of Stearoyl-CoA Desaturase 1 in Tumor Tissues Are Early Diagnostic Markers for Ovarian Cancer. *Cancer Res. Commun.* **3**, 1840–1852. <https://doi.org/10.1158/2767-9764.CRC-23-0138> (2023).

Author contributions

M. T-O., K. H., and Y. M. conceived the study, designed the experiments, analyzed and interpreted the data. Y. M. wrote the first draft of the manuscript. M. T-O. and Y. M. acquired the data. Takashi Yamamoto, T. M., K. I., K. O., K. N., Takafumi Yokota, M. M., K. Y., S. T., M. O., and Tomoyuki Yamasaki collected study material. M. T-O. reviewed and edited the manuscript. All authors reviewed the final manuscript.

Funding

This work was supported by Grant-in-Aid for Scientific Research (C) no. 23K06880 and 21K07409 from the Ministry of Education, Culture, Sports, Science and Technology of Japan, Research Grant from Nakatani Foundation, and Medical Research Grant Program (A), Osaka Foundation for the Prevention of Cancer and Cardiovascular Diseases.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-025-94496-y>.

Correspondence and requests for materials should be addressed to Y.M.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2025