

## Direct analysis of aberrant glycosylation on haptoglobin in patients with gastric cancer

Jae-Han Kim<sup>1</sup>, Sung Hyeon Lee<sup>2</sup>, Sookyung Choi<sup>2</sup>, Unyong Kim<sup>3</sup>, In Seok Yeo<sup>4</sup>, Su Hee Kim<sup>2</sup>, Myung Jin Oh<sup>3</sup>, Hantae Moon<sup>3</sup>, Jua Lee<sup>3</sup>, Seunghyup Jeong<sup>3</sup>, Min Gew Choi<sup>5</sup>, Jun Ho Lee<sup>5</sup>, Tae Sung Sohn<sup>5</sup>, Jae Moon Bae<sup>5</sup>, Sung Kim<sup>5</sup>, Yang Won Min<sup>5</sup>, Hyuk Lee<sup>5</sup>, Jun Haeng Lee<sup>5</sup>, Poong-Lyul Rhee<sup>5</sup>, Jae J. Kim<sup>5</sup>, Su Jin Lee<sup>5</sup>, Seung Tae Kim<sup>5</sup>, Jeeyun Lee<sup>5</sup>, Se Hoon Park<sup>5</sup>, Joon Oh Park<sup>5</sup>, Young Suk Park<sup>5</sup>, Ho Yeong Lim<sup>5</sup>, Won Ki Kang<sup>5</sup>, Hyun Joo An<sup>3</sup>, Jung Hoe Kim<sup>4</sup>

<sup>1</sup>Department of Food and Nutrition, Chungnam National University, Yuseong-Gu, Daejeon, Korea

<sup>2</sup>GLYCAN Co., Ltd., Healthcare Innovation Park, Bundang-Gu, Seongnam, Korea

<sup>3</sup>Graduate School of Analytical Science and Technology, Chungnam National University, Yuseong-Gu, Daejeon, Korea

<sup>4</sup>Department of Biological Sciences, Korea Advanced Institute of Science and Technology, Yuseong-Gu, Daejeon, Korea

<sup>5</sup>Samsung Medical Center, Sungkyunkwan University School of Medicine, Seoul, Korea

**Correspondence to:** Se Hoon Park, **email:** hematoma@skku.edu

**Keywords:** stomach neoplasm, haptoglobin, N-glycan, glycoprotein, biomarker

**Received:** October 05, 2016

**Accepted:** December 21, 2016

**Published:** December 29, 2016

### ABSTRACT

Based on our previous studies, differential analysis of N-glycan expression bound on serum haptoglobin reveals the quantitative variation on gastric cancer patients. In this prospective case-control study, we explore the clinically relevant glycan markers for gastric cancer diagnosis. Serum samples were collected from patients with gastric cancer ( $n = 44$ ) and healthy control ( $n = 44$ ). N-glycans alteration was monitored by intact analysis of Hp using liquid chromatography–mass spectrometry followed by immunoaffinity purification with the serum samples. Intensity and frequency markers were defined depending on the mass spectrometry data analysis. Multiple markers were found with high diagnostic efficacy. As intensity markers (*I*-marker), six markers were discovered with the AUC > 0.8. The high efficiency markers exhibited AUC of 0.93 with a specificity of 86% when the sensitivity was set to 95%. We additionally established frequency marker (*f*-marker) panels based on the tendency of high N-glycan expression. The AUC to conclude patients and control group were 0.82 and 0.79, respectively. This study suggested that N-glycan variation of serum haptoglobin were associated with patients with gastric cancer and might be a promising marker for the cancer screening.

### INTRODUCTION

Gastric cancer (GC) is the most frequently occurring malignancy in Korea, and one of the main causes of cancer death [1]. When patients are diagnosed early their 5-year survival rate is approximately 90% [2]. Unfortunately, the vast majority of GC patients present with advanced stage disease and the overall prognosis remains very poor. Although screening endoscopy and/or upper gastrointestinal series (UGIS) are widely performed in clinical settings to detect early-stage GC, to date no randomized study has been able to demonstrate a reduction in gastric cancer mortality [3].

Aberrant glycosylation is a major post-translational modification (PTM) of proteins, frequently observed in cancer cells and tissues [4–6]. In particular, a variety of cancers display changes in glycan structures of serum proteins such as IgG,  $\alpha$ -fetoprotein (AFP), prostate-specific antigen (PSA) and haptoglobin (Hp) [7, 8]. Hp, a major acute-phase glycoprotein comprising 0.4–2.6% of total blood proteins, consists of two  $\alpha$ - and two  $\beta$ -subunits whose glycosylation level changes in various types of cancer and inflammation [8, 9]. Hp is a highly sialylated glycoprotein with four N-glycosylation sites in  $\beta$ -chain at Asn 184, 207, 211, and 241 in  $\beta$ -chain [10]. It has been demonstrated that the glycosylated Hp is resulted directly

from cancer itself rather than secondary to cancer-induced inflammation [11]. In order to discover novel glycan-related biomarkers, mass spectrometry (MS) has been widely used because of the advanced features including high sensitivity, stoichiometry and dynamic range [12–15]. Analysis of serum N-glycan profiles may readily allow the early diagnosis of various type of cancers, occupying a niche on the non-invasive, *in vitro* diagnosis area [7]. Recently, MS analyses of intact glycoprotein have been actively employed to determine characterization of therapeutic glycoproteins as well as to discover biomarkers associated with disease status [16–18]. Instead of glycan profiling, direct analysis of intact glycoprotein has merits such as simple, easy handling of sample preparation and time saving for analysis. This method could have thereby high potential to discover the novel cancer biomarkers with aberrant glycosylation.

Previously we explored the glycosylation of serum Hp whether the glycan alteration can be clinically relevant biomarkers for GC screening [8]. N-glycans bound on serum Hp were released enzymatically, and then the expression profile of N-glycans was monitored by nanoLC Chip Q-TOF system. As a result, we observed clear and clinically significant deviations between cancer patients and control groups suggesting the potential glycan biomarkers for GC diagnosis. Based on these considerations, we pursued the N-glycan alteration by analyzing intact Hp without releasing and collection of glycans from a protein. Beyond a fine analytical instrument, MS has been evaluated as a diagnostic platform technology that is able to directly measure the aberration of a glycoprotein with glycan heterogeneity.

## RESULTS

Clinical characteristics of gastric cancer patients and healthy controls were summarized in Table 1. The study population has 44 cancer-free, healthy controls and 44 GC patients in a random manner. Forty-one GC patients had metastatic disease, and three patients gave their serum samples prior to removal of primary tumor. All study participants were Koreans.

Collected serum samples were first applied to immunoaffinity purification. No visual contamination was found on the SDS-PAGE after purification. In addition to serum samples obtained for the present study, commercially available serum ( $n = 7$ ) was additionally subjected in parallel to ensure the quality of sample preparation and sequential MS analysis. The yield of purified Hp detected by quantification assay was estimated to be 100–500  $\mu\text{g}$  per 450  $\mu\text{L}$  of serum. Among them, 20  $\mu\text{g}$  of purified serum Hp was used for the subsequent preparation of N-glycan for LC-MS analysis.

### Mass spectrometric monitoring of intact Hp

Previously we observed the significant variation in N-glycan expression on GC patients by MS profiling [8].

Though the aberrant glycosylation suggests a strong potential as a diagnostic biomarker, it requires the enzymatic release of N-glycan that necessitate a day of reaction and purification time. In the present study, glycosylation on Hp isolated by immunoprecipitation was directly monitored in a glycoprotein level without releasing N-glycan.

Because multiple charged ion species of Hp were produced by electrospray ionization, a complex mass spectrum carrying the information on the molecular weight corresponded to each glycoform on the  $\beta$ -chain of Hp was generated (Figure 1A). After the deconvolution of the raw spectrum, potential glycoforms of Hp could be identified. The overlay of deconvoluted mass spectra from healthy control and GC patient serum exhibits high visual reproducibility on mass accuracy but variations on peak intensities that suggest the possibility of potential markers (Figure 1B). Peaks of glycosylated Hp were assigned by the accurate masses and glycan correlations. N-glycans heterogeneity of and correlation through carbohydrate (monosaccharide) unit are biological characteristics of protein-bound glycans [19, 20]. Human N-glycans were biologically synthesized through the competitive but sequential transfer of four unit carbohydrates, hexose (Hex), N-acetylhexosamine (HexNAc), fucose (Fuc), and N-acetylneuraminic acid (NeuAc) [21–23]. From the common core structures, specific glycosyltransferases added these unit carbohydrates, therefore, each N-glycan bound to Hp should exhibit the unique mass deviation of those four unit carbohydrates showing multiple peak from a single protein [8, 24, 25].

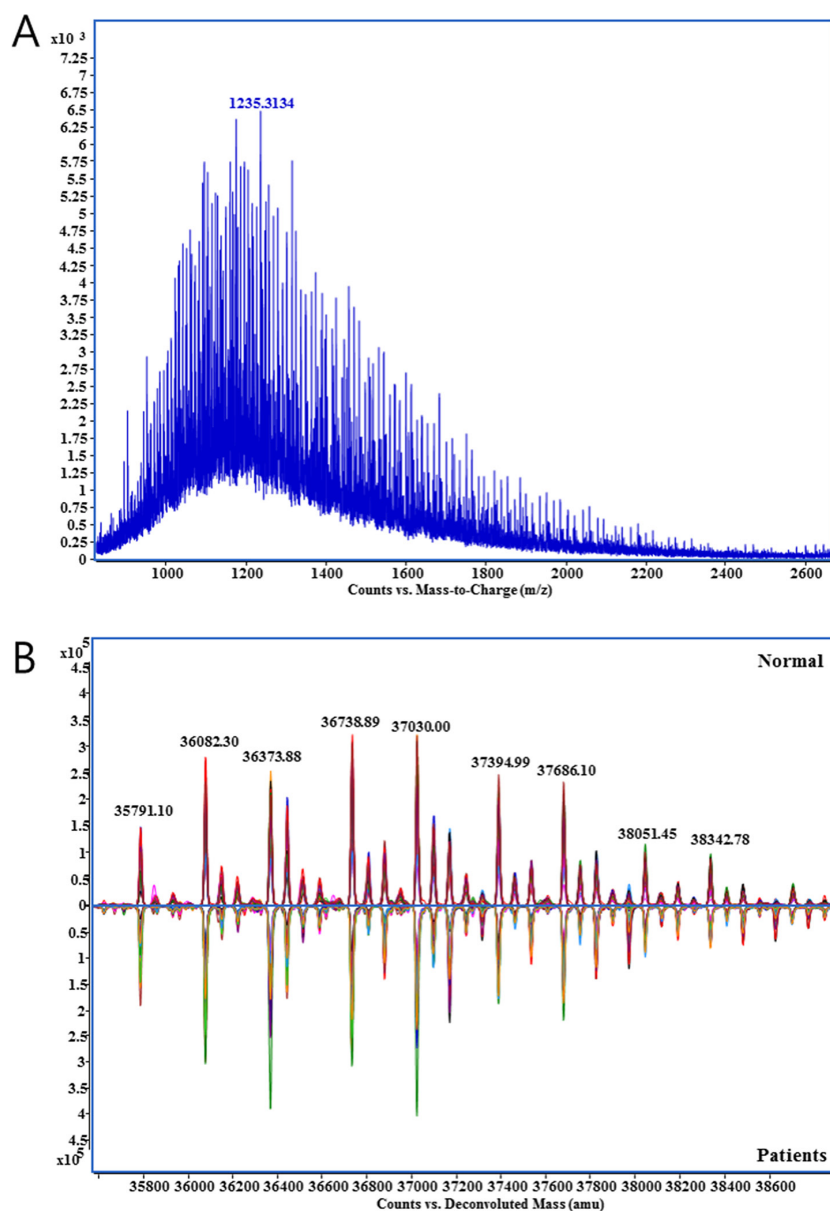
In a mass spectrum, a group of peaks having a correlation linked by the mass of each unit carbohydrate (Figure 2A) could be found. As described in Figure 2B, peak of  $m/z$  35791.10 and  $m/z$  36082.30 has the mass difference of NeuAc (291.25), then  $m/z$  36082.30 can be connected to  $m/z$  36373.88 and  $m/z$  36228.80 by NeuAc and Fuc unit, respectively. As such, it could be calculated that ion molecules of  $m/z$  36681.68 contain one Hex, one Fuc, and two NeuAc (glycan code-1:0:1:2) of additional glycan residues on the molecules of  $m/z$  35791.10. These sequential correlations clearly indicate that the signals are obtained from N-glycan variation of a single protein. From seven repetitive MS analyses with standard Hp from commercial serum, we observed 74 peaks that exhibit the single step correlation of unit carbohydrate.

### Analytical stability and reproducibility

Hp purified from a commercial serum was used as a quality control standard for clinical sample analysis. Each clinical sample was analyzed blindly and randomly repeated three times which made up of seven batches of sample run. More than one standard Hp from commercial sera were included in each sample batch to ensure the reproducibility and stability of analysis. As a result, MS profile of standard Hp showed extremely

**Table 1: Clinical characteristics of study participants**

	Control (N = 44)	Gastric cancer (N = 44)
Age, years		
Median	48	55
Range	28 to 65	41 to 70
Gender		
Male	11	28
Female	33	16
Tumor stage		
1-3		3
4		41



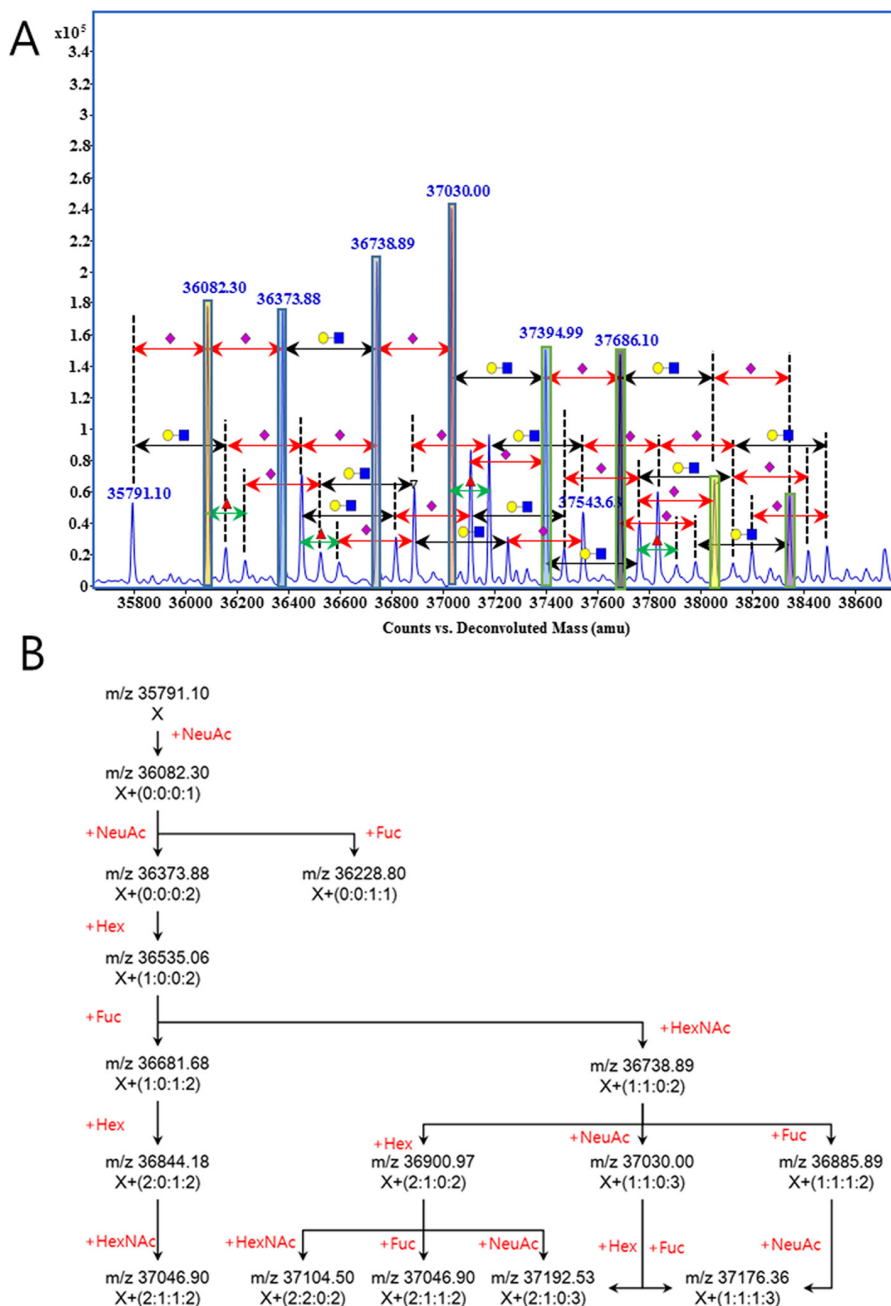
**Figure 1: Mass spectrum obtained from the direct analysis of haptoglobin.** (A) Raw mass spectrum with multiply charged haptoglobin glycoforms., (B) Deconvoluted spectrum of haptoglobin glycoforms from healthy control ( $n = 44$ , upper) and gastric cancer patient group ( $n = 44$ , lower). Spectra in each groups were overlaid.

high reproducibility, with the range of 0.940–0.998 of a *Pearson* correlation coefficient through the whole process. Moreover, the average coefficient of variation (CV) of peak intensity in a sample was maintained under 18% indicating high quantitative reproducibility of this entire analysis process (Table 2). Daily variation of MS analyses was in acceptable ranges as well. MS profile of standard Hp analyzed in a same day showed the high reproducibility of  $r^2 > 0.99$  (Figure 3A). Mass profile of standard Hp

between different days of operation exhibits slight less yet still high reproducibility of  $r^2 > 0.94$  (Figure 3B).

### Discovery of screening makers for gastric cancer

After collectively integrated each group's MS data, every peaks eventually have two data domains, Intensity and Frequency. Intensity represents an ion count of peak presented in each sample. Each peak has intensity values as



**Figure 2: A set of glycan correlation observed in the mass spectrometric analysis of human haptoglobin.** Glycan correlation was displayed on the mass spectrum (A), and illustrated as schematic diagram (B). Symbols in figure 2(A) were ●, hexose; ■, N-acetylhexosamine; ▲, fucose; and ◆, N-acetylneuraminic acid. Number in parenthesis on figure 2(B), are the number of carbohydrate units attached on the molecule X ( $m/z$  35791.10). Number and type of units were calculated from the accurate mass and indicates hexose, N-acetylhexosamine, fucose and N-acetylneuraminic acid from left to right, respectively.

**Table 2: Average CV of peak intensities obtained seven independent analysis of haptoglobin isolated from commercial human sera**

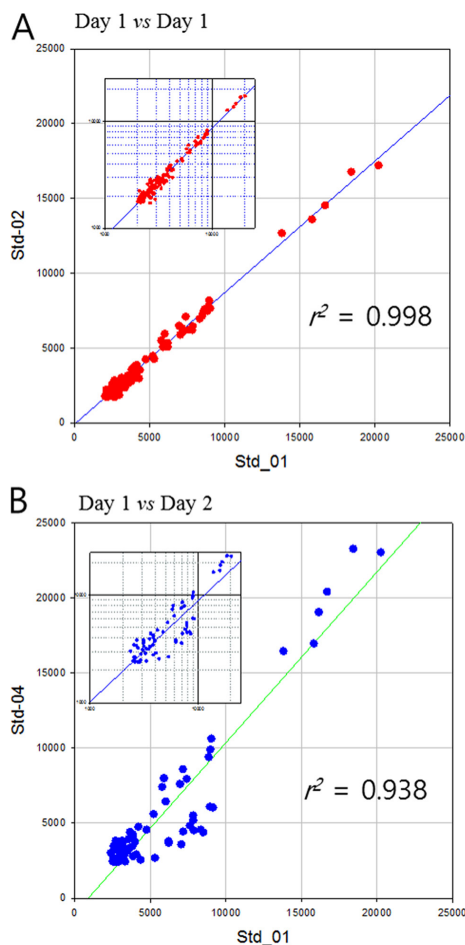
Frequency of peak	Average CV
100~80%	18.4%
80~60%	15.9%
60~40%	15.0%

much of its observations throughout the analysis of patient and control groups. Meanwhile, frequency (presence) of a peak indicates percentile of the sample that contains the peak among total samples resulting a single value per sample groups. Markers were found independently from these two data domains designated intensity marker (*I*-marker) and frequency marker (*f*-marker).

*I*-marker is a peak that exhibits the significant difference of their intensities between patient and healthy control groups, which is traditionally used in the diagnosis of diseases. To avoid the case of ‘N/A’ or ‘unable to diagnose’, the marker peak should show up more than

80% frequency (*f*) in both sample groups ( $f(\text{patient}) > 80\%$  and  $f(\text{control}) > 80\%$ ). Similarly, *f*-marker is an integrative marker that could be deduced from a group of peaks. Regardless of their intensity, *f*-marker only counts the tendency (occurrence) of their presence in patient group or healthy control group.

To discover the *I*-markers, student’s *t*-test was performed initially with the consideration of peak frequency searching potential marker candidates. Among the peaks with  $P < 10^{-5}$ , the diagnostic markers were selected only if the AUC of ROC curve was higher than 0.80. Six screening markers were determined finally with



**Figure 3: Pearson’s correlation of mass spectrometric dataset of standard haptoglobin samples. (A) analyzed same day and (B) different days. Small graphs are log-log plot to clarify the dispersion of small values.**

the AUC range 0.81 to 0.93 (Table 3, Figure 4A–4F). The best predictive marker for gastric cancer was  $m/z$  37978.98 corresponded to Hex<sub>22</sub>HexNAc<sub>18</sub>Fuc<sub>2</sub>NeuAc<sub>10</sub> which showed AUC 0.93 (Figure 4G). Though the sensitivity and specificity of marker are moving along with the ROC curve line depending on the threshold values, it showed the 85% sensitivity when the specificity was fixed at 96%.

*f*-markers were defined as a group of N-glycans that expressed more frequently in either patient or healthy control group. Regardless of their intensity, peaks with the ratio of frequency above 1.5 fold change were collected and categorized in two groups depending on their tendency. If a group of N-glycans observed more frequently in patient group, we called it a positive *f*-marker. If a group of N-glycans had opposite trait, it was designated as negative *f*-marker. Then, a score was estimated by counting the number of peaks presence in the sample. As a result, each sample could have a positive *f*-marker value and another negative *f*-marker value. The average score (count) of positive *f*-marker in patient group was  $5.4 \pm 2.7$ , while in control group was  $2.3 \pm 1.8$ . Four negative *f*-marker, average score of control group was  $7.7 \pm 3.1$  which was high enough to segregate the patient group having a value of  $3.6 \pm 3.4$ . The *P*-value and the AUC of ROC of positive *f*-marker were  $7.81 \times 10^{-9}$  and 0.82, respectively. Negative *f*-marker which was able to positively distinguish healthy control group exhibited the *P*-value of  $6.3 \times 10^{-7}$  and AUC of 0.79 (Table 4, Figure 5).

## DISCUSSION

When it comes to GC, a significant geographic variation in the incidence exists, with the highest rates being reported in East Asian countries [26], including Korea, China and Japan. In Korea and Japan, where mass screening is performed widely, early detection is often possible [27]. However, mass screening for GC, particularly by endoscopy, may not be the most practical approach because of reasons such as acceptance, availability, and cost [3]. Even in Korea where a nationwide screening program provides free-of-charge biennial endoscopy for men and women aged 40 or over [28], only 43% of the targeted population had undergone GC screening [29]. Besides gastric endoscopy, noninvasive methods for early detection of GC include UGIS and serum pepsinogen assay [29]. These screening programs have meant that patients are more likely to be diagnosed with early GC so that have it treated effectively. Unfortunately, the sensitivity of UGIS and serum pepsinogen has been found to be less than 50% in detecting early GC [29], and it is still unclear whether the screening reduces the number of people who eventually die from the disease [3].

Although the use of tumor biomarkers has been proposed for decades, the discovery of specific genetic or protein biomarkers has been fundamentally

complex due to technical issues with comprehensive expression platforms, limitations in multiplex clinical assay development, and most importantly an incomplete understanding of tumor biology. The use of accurate serum biomarkers to evaluate patients with GC could, if found, lead not only to early detection but also to potential reduction in healthcare costs. Recently, a growing body of evidence supports crucial roles for glycans at a number of pathophysiological steps of tumor progression [30]. Glycans are covalent assemblies of sugars that exist in either free-form or in covalent complex with proteins or lipids, and surfaces of all mammalian cells are covered with a thick layer of complex glycans [31]. Importantly, the altered glycosylation is one of the most common and most complex PTM and plays a critical role in cancer cell biology hallmarks such as proliferative signaling, immortality, angiogenesis, invasion and metastasis [32].

Interestingly, serum measurement of certain glycans is already used to facilitate diagnosis, surveillance, or provide a surrogate measure for therapeutic outcome. For example, although the sensitivity and specificity were not very high enough to gain grade 1 recommendations [33], the serological markers CA125, CA19-9, AFP and PSA are biomolecules decorated with glycans that are commonly overexpressed by a number of cancer types. The majority of serum proteins are glycosylated, and changes in protein glycosylation are recently important focus in biomarker studies. In particular, serum Hp, an acute-phase glycoprotein, is highly expressed in a variety of cancer types [34], but its glycosylation status is different from one type of cancer to another [35]. Previous reports demonstrated that original tissue and cell producing aberrant glycosylation of serum Hp is different in each cancer type [9, 11, 36, 37]. In GC, a previous study showed the increased level of SLe<sup>x</sup> in serum Hp based on western blotting analysis with anti-SLe<sup>x</sup> antibody [38]. Furthermore, our previous study showed that aberrant glycan structures of fucosylation linked to SLe<sup>x</sup> epitope in GC that was revealed by N-glycan profiling base on LC-MS analysis [8].

Technical advances in analytical methodology have greatly enhanced our capabilities for analysis of human glycans, although inherent diversity and complexity make analytic process extremely difficult. Especially, MS is emerging as a powerful technique for glycan profiling and structural elucidation, and can be used as a precise tool to facilitate identification of aberrant glycosylation related to cancers [12, 39, 40]. It was found that, by MS analysis, GC patients have distinctive serum N-glycan patterns including the high-mannose-type glycans, glycans with one complex type antenna, bigalactosylated biantennary glycans, and nongalactosylated biantennary glycans [41]. However, most of previous studies utilized glycan itself released from carrier molecules (*e.g.*, protein, lipid, other organic compound) for MS analysis, a time consuming process to release glycans in sample preparation because it needs

**Table 3: Clinical efficacy of intensity markers (*I*-marker)**

	<i>m/z</i>	Patient				Healthy control				<i>p</i> -value	AUC
		AVE	±	SD	CV	AVE	±	SD	CV		
Marker A	36593.422	3899	±	1257	32%	2535	±	679	27%	3.2.E-05	0.83
Marker B	36885.354	9615	±	4027	42%	5512	±	1693	31%	1.4.E-07	0.81
Marker C	37176.968	11483	±	5601	49%	5977	±	1864	31%	4.0.E-08	0.81
Marker D	37248.766	4293	±	1358	32%	2802	±	733	26%	3.6.E-07	0.84
Marker E	37541.054	7192	±	2756	38%	4048	±	1227	30%	6.6.E-09	0.83
Marker F	37978.984	6175	±	3049	49%	2585	±	598	23%	3.3.E-07	0.93

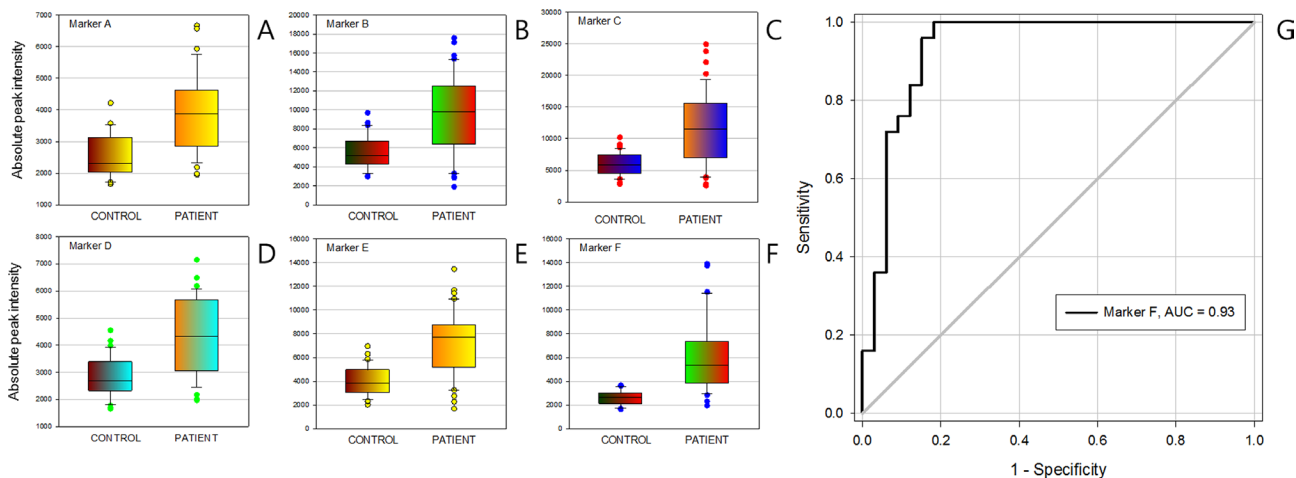
*m/z*, molecular mass; AVE, average; SD, standard deviation; CV, coefficient of variation; AUC, area under the curve.

an enzymatic reaction overnight. In the present study, we applied a MS technique focused on intact molecular weight of specific glycoprotein to solve the problems. It is not necessary to release glycans from human biomolecules in case of the MS analysis for intact glycoprotein, which makes the method to be able to use clinically with several benefits such as simple, easy handling of sample preparation and time saving for entire analysis.

In the present case-control study, we attempted to explore differential expressions of Hp N-glycan in GC by intact MS analysis, and to evaluate the diagnostic performance of N-glycan variation of intact Hp as a potential screening marker for GC. To identify candidate glycan markers on intact Hp, we initially focused on intact *m/z* signals which were quantitatively strong enough to distinguish the patients and healthy control groups and whose isotopic peak distribution was apparent in all samples. The diagnostic efficacy based on *P*-values and AUC was evaluated in 88 serum samples (44 healthy controls and 44 patients) that had been randomly selected. Conclusively, we found more than a few aberrant glycan markers bound to serum Hp that have high sensitivity as well as high specificity in differential diagnosis of GC patients and healthy controls. One may argue that the accuracy of a

screening test is indicated by its sensitivity and specificity. Our novel glycan biomarkers have high sensitivity as well as specificity. Furthermore, when combined together, the AUC of the glycan biomarkers reached as high as AUC 0.8 ~ 0.93. Because GC is considered a heterogeneous disease, it seems unlikely that one biomarker will be uniformly elevated. Therefore, a panel of biomarkers (i.e., multi-biomarker) may provide more accurate diagnosis than any given marker alone. To the authors' best knowledge, this is the first and only study to prove a diagnostic value of aberrant glycans bound to serum Hp based on intact molecular weighting in cancer patients.

Our findings have limitations, and warrant further validation studies including in a larger study as well as in patients with other cancer types than GC. Our results showed a disparity in baseline characteristics between GC patients and control, although studies suggested glycomic profiling may be dependent on the subjects' age and gender [42]. A majority of GC patients had an advanced stage disease. The small sample size is also inadequate to make any definitive conclusions on the performance of our finding, so the validation study involving a larger number of samples is warranted. In addition, because the sample lacked diversity with respect to ethnic group, it is

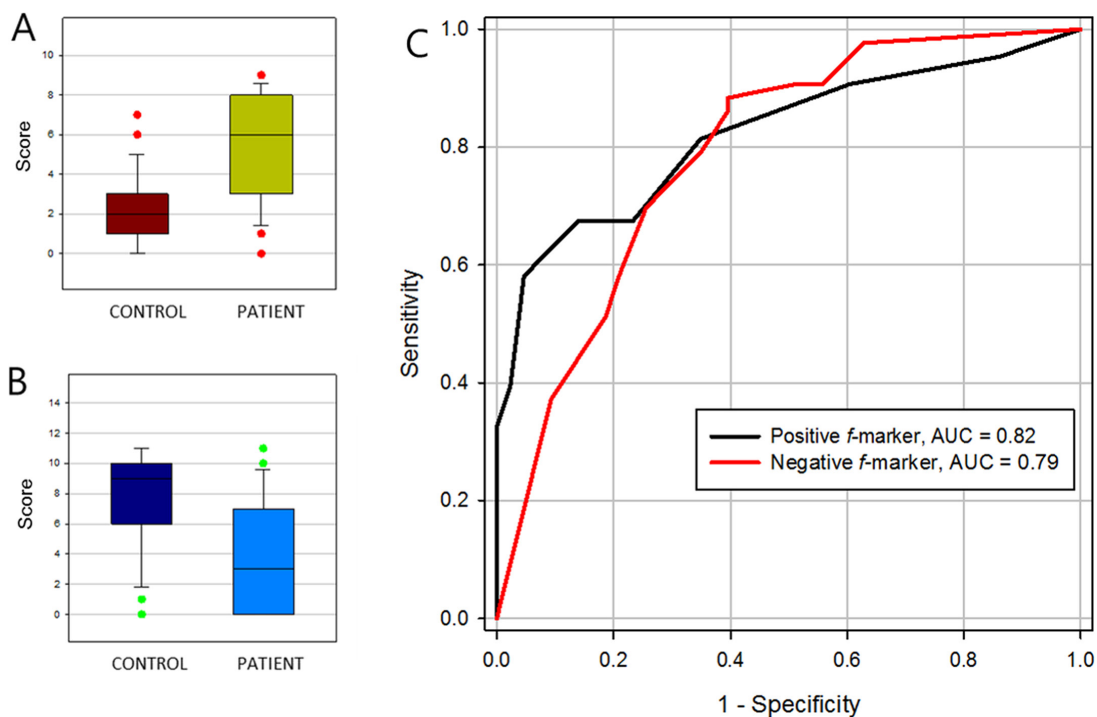


**Figure 4: Clinical efficacy of intensity markers (*I*-marker).** (A) to (F) were Box-plot of intensity marker A to marker F, (G) the receiver operating characteristic curve (ROC) of intensity marker (F).

**Table 4: Clinical efficacy of frequency marker (*f*-marker)**

	Frequency		Frequency ratio	Count						P-value	AUC
	f(N)	f(P)		Control group			Patient group				
Positive <i>f</i> -marker	26%	72%	2.8								
	49%	88%	1.8								
	23%	63%	2.7								
	49%	77%	1.6								
	14%	56%	4.0	2.3	±	1.8	5.4	±	2.7	7.81.E-09	0.82
	33%	70%	2.1								
	7%	37%	5.3								
	9%	33%	3.5								
	16%	47%	2.9								
Negative <i>f</i> -marker	79%	26%	3.1								
	74%	44%	1.7								
	63%	26%	2.5								
	40%	7%	5.7								
	86%	42%	2.1								
	56%	28%	2.0	7.7	±	3.1	3.6	±	3.8	6.31.E-07	0.79
	86%	53%	1.6								
	91%	53%	1.7								
	60%	30%	2.0								
	60%	23%	2.6								
72%	30%	2.4									

f(N), frequency in control group; f(P), frequency in gastric cancer patient group; Count, numbers of *f*-markers in each sample; AUC, area under the curve.



**Figure 5: Clinical efficacy of frequency markers (*f*-marker).** (A) positive *f*-marker, (B) negative *f*-marker and (C) their receiver operating characteristic (ROC) curve.



unclear the finding should also be applied to other race than Korean. When it comes to the prognosis of patients with GC, although randomized clinical trials showed similar magnitude of survival benefit with chemotherapy [43], it is known that survival outcome and tumor biology may differ considerably between Western and Asian countries. However, the results of the present study provide evidence for serum measurements for aberrant glycans bound to serum Hp in early detection of GC, although the mechanism underlying the alterations in glycans has not been fully understood yet. Moreover, the role of these aberrant glycosylations in relation to prognosis in GC is outside the scope of this study, and remains as a topic for extensive further studies. Further studies focused on glycan alterations in cancer cells or tissues may help to identify cancer-specific glycan pathways. In addition to the role of glycans as cancer biomarkers [44], the broad spectrum of glycans that activate cancer proliferation, invasion and metastasis may lead to a range of glycan-specific targeting approaches for novel drug development [45]. Our results provide further insight into the glycomics based on MS screening in cancer development, and may provoke more detailed investigations leading to identification of a panel of diagnostic serological biomarkers applicable to early detection of cancer.

## CONCLUSIONS

We studied whether N-glycan variation in serum Hp based on intact *m/z* signals is occurred and showed high efficiency diagnostic values as screening markers for GC. Despite a small study with only 88 samples tested, we might conclude that aberrant glycans bound to serum Hp are associated with patients with gastric cancer and are a potential novel marker for gastric cancer screening. Further, multiple biomarkers with high diagnostic efficacy might enable the biomarker panel test, results of which can be obtained from a single MS analysis.

## MATERIALS AND METHODS

This was a prospective case-control study, in which a total of 88 serum samples were obtained from 44 patients with GC and from 44 cancer-free healthy volunteers (Table 1) at Samsung Medical Center (SMC, Seoul, Korea). All serum samples were collected and processed in an identical fashion and the cancer and control samples were obtained over the same study periods. This study protocol (ClinicalTrials.gov identifier: NCT02848066) was reviewed and approved by the SMC Institutional Review Board (IRB# SMC2015-07-146-001). All participating subjects provided written informed consent in accordance with the Declaration of Helsinki, after having been informed about the purpose and investigational nature of the study.

## Immunoaffinity purification of serum Hp

Serum Hp purification was performed using Hp antibody-conjugated affinity column [11]. In brief, 450  $\mu$ L serum from each subject was diluted in 4.5 mL phosphate-buffered saline (PBS, pH 7.4), and then applied to the anti-Hp affinity column. After washing the unbound components with 30 mL PBS and elution of Hp with elution buffer (0.1 M glycine, 0.5 M NaCl, pH 2.8), and the eluent was fractionated into a tube containing neutralization buffer (1.0 M Tris-HCl, pH 9.0). The detergent was removed and eluent samples were concentrated using a centrifugal filter (MWCO 10,000, Amicon Ultra, Millipore; Billerica, MA, USA). Quantification assay of purified Hp was performed using a Quant-iT Assay Kit (Invitrogen; Carlsbad, CA, USA). To check the purity of Hp, the eluent was randomly applied to 12.5% SDS-PAGE with Coomassie Brilliant Blue staining. Each sample was lyophilized and kept at  $-80^{\circ}\text{C}$  until analysis.

## Denaturation and alkylation of Hp

Twenty micrograms of purified Hp were dissolved in a buffer, which consisted of 50 mM  $\text{NH}_4\text{HCO}_3$  (pH 9.0), 5 mM DTT, and 14 mM IAA. The Hp dissolved in buffer solution was boiled in water for 10 min at  $90^{\circ}\text{C}$  to reduce the disulfide bridges in  $\alpha$ - and  $\beta$ - subunits of Hp, and reduced sulfhydryl groups were alkylated with IAA to prevent reconstruction of disulfide bridges. Finally, denatured Hp was evaporated in vacuum, and then reconstituted with 200  $\mu$ L of distilled water. One microliter (100 ng of Hp) was injected into the nanoLC-chip/Q-TOF MS system.

## NanoLC-chip/Q-TOF MS analysis of Hp

The analytical chip (Agilent Technologies Inc., San Jose, CA, USA) consisted of  $9 \times 0.075$  mm enrichment column and  $43 \times 0.075$  mm analytical column both of which were packed with 5  $\mu$ m C8-bonded particles. A rapid elution gradient for Hp was delivered to the analytical column at 300 nL/min using mobile phases of (A) 0.1% formic acid (v/v) in water, and (B) 0.1% formic acid (v/v) in acetonitrile, ramping from 3% to 25% B over 1 min, from 25% to 60% B over 3 min, from 60% to 80% B over 4 min, and from 80% to 95% B over 4 min. The columns were flushed out with 95% B prior to re-equilibration.

Following chromatographic separation, Hp was ionized by chip-integrated nano-ESI spray tip and analyzed by a Q-TOF mass analyzer (model 6540, Agilent Technologies Inc., San Jose, CA, USA). MS spectra were acquired in positive ionization mode over mass range of *m/z* 500 to 4500 with a scan rate of 1.5 spectra/sec. After data acquisition, raw LC/MS data was deconvoluted by maximum entropy algorithm using MassHunter software equipped with BioConfirm add-on (Agilent Technologies Inc., San Jose, CA, USA).

## Statistical analysis

Glycan correlation between peaks was calculated based on the average mass from seven independent analyses of Hp. Unit mass of hexose (162.14), N-acetylhexosamine, (203.19), fucose (146.14), and N-acetylneuraminic acid (291.25) was used. Mass tolerance for the assignment of carbohydrate correlation was less than 25 ppm. Student *t*-test was used for comparison between patient and healthy control group to screen potential markers with the assumption of two-tailed and equal variance. Area under the curve (AUC) of receiver operating characteristic (ROC) curves was calculated by the trapezoidal rule. Data handling, descriptive and statistical analyses were carried out by Microsoft Excel 2013 (Microsoft, Seattle, WA, USA), SigmaPlot 12.5 (Systat Software Inc., San Jose, CA, USA) and Stata 12 (Stata Corp., College station, TX, USA).

## ACKNOWLEDGMENTS AND FUNDING

This project was supported by Glycan Co., Ltd (Seongnam, Republic of Korea).

## CONFLICTS OF INTEREST

Sung Hyeon Lee, Sookyung Choi and Su Hee Kim are employees of Glycan Co., Ltd (Healthcare Innovation Park, 172 Dolma-ro, Bundang-gu, Seongnam 13605, Korea). No potential conflict of interest is declared by other authors.

## REFERENCES

1. Jung KW, Won YJ, Kong HJ, Oh CM, Cho H, Lee DH, Lee KH. Cancer statistics in Korea: incidence, mortality, survival, and prevalence in 2012. *Cancer Res Treat.* 2015; 47:127–141.
2. Park SH, Kim DY, Heo JS, Lim DH, Park CK, Lee KW, Choi SH, Sohn TS, Kim S, Noh JH, Kim YI, Park JO, Kim K, et al. Postoperative chemoradiotherapy for gastric cancer. *Ann Oncol.* 2003; 14:1373–1377.
3. Leung WK, Wu MS, Kakugawa Y, Kim JJ, Yeoh KG, Goh KL, Wu KC, Wu DC, Sollano J, Kachintorn U, Gotoda T, Lin JT, You WC, et al. Screening for gastric cancer in Asia: current evidence and practice. *Lancet Oncol.* 2008; 9:279–287.
4. An HJ, Gip P, Kim J, Wu S, Park KW, McVaugh CT, Schaffer DV, Bertozzi CR, Lebrilla CB. Extensive determination of glycan heterogeneity reveals an unusual abundance of high mannose glycans in enriched plasma membranes of human embryonic stem cells. *Mol Cell Proteomics.* 2012; 11:M111 010660.
5. Hua S, Saunders M, Dimapasoc LM, Jeong SH, Kim BJ, Kim S, So M, Lee KS, Kim JH, Lam KS, Lebrilla CB, An HJ. Differentiation of cancer cell origin and molecular subtype by plasma membrane N-glycan profiling. *J Proteome Res.* 2014; 13:961–968.
6. Hakomori S. Glycosylation defining cancer malignancy: New wine in an old bottle. *Proceedings of the National Academy of Sciences of the United States of America.* 2002; 99:10231–10233.
7. Adamczyk B, Tharmalingam T, Rudd PM. Glycans as cancer biomarkers. *Biochim Biophys Acta.* 2012; 1820:1347–1353.
8. Lee SH, Jeong S, Lee J, Yeo Is, Oh MJ, Kim U, Kim S, Kim SH, Park SY, Kim JH, Park SH, Kim JH, An HJ. Glycomic profiling of targeted serum haptoglobin for gastric cancer using nano LC/MS and LC/MS/MS. *Mol BioSyst.* 2016; 12:3611–3621.
9. Zhang S, Shang S, Li W, Qin X, Liu Y. Insights on N-glycosylation of human haptoglobin and its association with cancers. *Glycobiology.* 2016 (in press; doi:10.1037/C6MB00559D).
10. Ratanasopa K, Chakane S, Ilyas M, Nantasenamat C, Bulow L. Trapping of human hemoglobin by haptoglobin: molecular mechanisms and clinical applications. *Antioxid Redox Signal.* 2013; 18:2364–2374.
11. Park SY, Lee SH, Kawasaki N, Itoh S, Kang K, Hee Ryu S, Hashii N, Kim JM, Kim JY, Hoe Kim J. alpha1-3/4 fucosylation at Asn 241 of beta-haptoglobin is a novel marker for colon cancer: a combinatorial approach for development of glycan biomarkers. *Int J Cancer.* 2012; 130:2366–2376.
12. Lebrilla CB, An HJ. The prospects of glycan biomarkers for the diagnosis of diseases. *Mol Biosyst.* 2009; 5:17–20.
13. Ruhaak LR, Miyamoto S, Lebrilla CB. Developments in the identification of glycan biomarkers for the detection of cancer. *Mol Cell Proteomics.* 2013; 12:846–855.
14. Hua S, An HJ. Glycoscience aids in biomarker discovery. *BMB Rep.* 2012; 45:323–330.
15. An HJ, Kronewitter SR, de Leoz ML, Lebrilla CB. Glycomics and disease markers. *Curr Opin Chem Biol.* 2009; 13:601–607.
16. Rosati S, van den Bremer ET, Schuurman J, Parren PW, Kamerling JP, Heck AJ. In-depth qualitative and quantitative analysis of composite glycosylation profiles and other micro-heterogeneity on intact monoclonal antibodies by high-resolution native mass spectrometry using a modified Orbitrap. *MAbs.* 2013; 5:917–924.
17. Van Scherpenzeel M, Willems E, Lefeber DJ. Clinical diagnostics and therapy monitoring in the congenital disorders of glycosylation. *Glycoconj J.* 2016; 33:345–358.
18. Ongay S, Martin-Alvarez PJ, Neuss C and de Frutos M. Statistical evaluation of CZE-UV and CZE-ESI-MS data of intact alpha-1-acid glycoprotein isoforms for their use as potential biomarkers in bladder cancer. *Electrophoresis.* 2010; 31:3314–3325.
19. Geyer H, Geyer R. Strategies for analysis of glycoprotein glycosylation. *Biochim Biophys Acta.* 2006; 1764:1853–1869.

20. Balaguer E, Neuss C. Glycoprotein characterization combining intact protein and glycan analysis by capillary electrophoresis-electrospray ionization-mass spectrometry. *Anal Chem.* 2006; 78:5384–5393.
21. Hua S, An HJ, Ozcan S, Ro GS, Soares S, DeVere-White R, Lebrilla CB. Comprehensive native glycan profiling with isomer separation and quantitation for the discovery of cancer biomarkers. *Analyst.* 2011; 136:3663–3671.
22. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. *Nature Reviews Molecular Cell Biology.* 2012; 13:448–462.
23. Aldredge D, An HJ, Tang N, Waddell K, Lebrilla CB. Annotation of a serum N-glycan library for rapid identification of structures. *J Proteome Res.* 2012; 11:1958–1968.
24. Ivatt RJ. Regulation of glycoprotein biosynthesis by formation of specific glycosyltransferase complexes. *Proc Natl Acad Sci USA.* 1981; 78:4021–4025.
25. Chandler KB, Pompach P, Goldman R, Edwards N. Exploring site-specific N-glycosylation microheterogeneity of haptoglobin using glycopeptide CID tandem mass spectra and glycan database search. *J Proteome Res.* 2013; 12:3652–3666.
26. Macdonald JS. Gastric cancer: Nagoya is not New York. *J Clin Oncol.* 2011; 29:4348–4350.
27. Kim YG, Kong SH, Oh SY, Lee KG, Suh YS, Yang JY, Choi J, Kim SG, Kim JS, Kim WH, Lee HJ, Yang HK. Effects of screening on gastric cancer management: comparative analysis of the results in 2006 and in 2011. *J Gastric Cancer.* 2014; 14:129–134.
28. Choi KS, Jun JK, Suh M, Park B, Noh DK, Song SH, Jung KW, Lee HY, Choi IJ, Park EC. Effect of endoscopy screening on stage at gastric cancer diagnosis: results of the National Cancer Screening Programme in Korea. *Br J Cancer.* 2015; 112:608–612.
29. Choi KS, Kwak MS, Lee HY, Jun JK, Hahm MI, Park EC. Screening for gastric cancer in Korea: population-based preferences for endoscopy versus upper gastrointestinal series. *Cancer Epidemiol Biomarkers Prev.* 2009; 18:1390–1398.
30. Taniguchi N. Toward cancer biomarker discovery using the glycomics approach. *Proteomics.* 2008; 8:3205–3208.
31. Lauc G, Kristic J, Zoldos V. Glycans-the third revolution in evolution. *Front Genet.* 2014; 5:145.
32. Hakomori SI, Cummings RD. Glycosylation effects on cancer development. *Glycoconj J.* 2012; 29:565–566.
33. Molina R, Barak V, van Dalen A, Duffy MJ, Einarsson R, Gion M, Goike H, Lamerz R, Nap M, Soletormos G, Stieber P. Tumor markers in breast cancer- European Group on Tumor Markers recommendations. *Tumour Biol.* 2005; 26:281–293.
34. Langlois MR, Delanghe JR. Biological and clinical significance of haptoglobin polymorphism in humans. *Clin Chem.* 1996; 42:1589–1600.
35. Goodarzi MT, Turner GA. Reproducible and sensitive determination of charged oligosaccharides from haptoglobin by PNGase F digestion and HPAEC/PAD analysis: glycan composition varies with disease. *Glycoconj J.* 1998; 15:469–475.
36. Fujita K, Shimomura M, Uemura M, Nakata W, Sato M, Nagahara A, Nakai Y, Takamatsu S, Miyoshi E, Nonomura N. Serum fucosylated haptoglobin as a novel prognostic biomarker predicting high-Gleason prostate cancer. *Prostate.* 2014; 74:1052–1058.
37. Okuyama N, Ide Y, Nakano M, Nakagawa T, Yamanaka K, Moriwaki K, Murata K, Ohigashi H, Yokoyama S, Eguchi H, Ishikawa O, Ito T, Kato M, et al. Fucosylated haptoglobin is a novel marker for pancreatic cancer: a detailed analysis of the oligosaccharide structure and a possible mechanism for fucosylation. *Int J Cancer.* 2006; 118:2803–2808.
38. Bones J, Byrne JC, O'Donoghue N, McManus C, Scaife C, Boissin H, Nastase A, Rudd PM. Glycomic and glycoproteomic analysis of serum from patients with stomach cancer reveals potential markers arising from host defense response mechanisms. *J Proteome Res.* 2011; 10:1246–1265.
39. Liu H, Zhang N, Wan D, Cui M, Liu Z, Liu S. Mass spectrometry-based analysis of glycoproteins and its clinical applications in cancer biomarker discovery. *Clin Proteomics.* 2014; 11:14.
40. Hua S, Lebrilla C, An HJ. Application of nano-LC-based glycomics towards biomarker discovery. *Bioanalysis.* 2011; 3:2573–2585.
41. Ozcan S, Barkauskas DA, Renee Ruhaak L, Torres J, Cooke CL, An HJ, Hua S, Williams CC, Dimapasoc LM, Han Kim J, Camorlinga-Ponce M, Rocke D, Lebrilla CB, et al. Serum glycan signatures of gastric cancer. *Cancer Prev Res (Phila).* 2014; 7:226–235.
42. Ding N, Nie H, Sun X, Sun W, Qu Y, Liu X, Yao Y, Liang X, Chen CC, Li Y. Human serum N-glycan profiles are age and sex dependent. *Age Ageing.* 2011; 40:568–575.
43. Janowitz T, Thuss-Patientce P, Marshall A, Kang JH, Connell C, Cook N, Dunn J, Park SH, Ford H. Chemotherapy vs supportive care alone for relapsed gastric, gastroesophageal junction, and oesophageal adenocarcinoma: a meta-analysis of patient-level data. *Br J Cancer.* 2016; 114:381–387.
44. Dube DH, Bertozzi CR. Glycans in cancer and inflammation—potential for therapeutics and diagnostics. *Nat Rev Drug Discov.* 2005; 4:477–488.
45. Fuster MM, Esko JD. The sweet and sour of cancer: glycans as novel therapeutic targets. *Nat Rev Cancer.* 2005; 5:526–542.