The 6th International Conference on Biochemistry and Molecular Biology

Utilization of lectin capture strategy combined with label-free quantitative proteomics analysis reveals differential expression levels of glycoproteins from metastatic colorectal cancer patients

Juthamard Chantaraamporn[1], Daranee Chokchaichamnankit[2], Naiyarat Prasongsook[3], Chantragan Srisomsap[2], Jisnuson Svasti[1,2], Vorarattt Champattanachai[1,2,*]

[1]Appied Biological Science Program, Chulabhorn Graduate Institute
[2]Laboratory of Biochemistry, Chulabhorn Research Institute,Bangkok,Thailand
[3]Division of Medical Oncology, Phramongkutklao Hospital, Bangkok, Thailand
*E-mail: vorararatt@cri.or.th

Abstract

Glycosylation is an important post-translational modification of proteins. Several reports indicate that altered glycosylation involves in pathobiological processes of diseases, including cancer i.e., colorectal cancer (CRC) which is one of the top three common cancers worldwide. Here, we aim to identify and quantify serum glycoproteins of patients with metastatic CRC and healthy controls. Glycoproteins from pooled serum of metastatic CRC patients and healthy controls were enriched by wheat germ agglutinin (WGA), a lectin binding protein that specifically binds to GlcNAc/sialic acid residues. Subsequently, enriched WGA glycoproteins were insolution-digested with trypsin. The digested peptides were subjected into LC-MS/MS and label-free quantitative proteomic approaches were applied to compare the expression levels of glycoproteins between two groups. Twenty-four glycoproteins were identified and compared. Three glycoproteins were up-regulated whereas five glycoproteins were down-regulated in CRC patients compared to those in controls (relative ratio > ±1.5, p value < 0.05). In addition, cellular components, molecular functions and biological processes of the identified proteins were assessed by Protein Analysis Through Evolutionary Relationships (PANTER). Glycoproteins displayed the difference in their expression levels between two groups were further validated by immunoblot and enzyme-linked immunosorbent assays. This study may provide a new way to find potential biomarkers for CRC.

Introduction

Colorectal cancer (CRC) has been considered as health impacts worldwide. Based on the cancer incidence report from GLOBOCAN by World Health Organization, 2012, it ranks the second most common cancer in females and the third common cancer in males.[1] CRC is preventive and curable if detected at early stages before the development of metastasis. There are various screening methods to detect CRC, such as colonoscopy and sigmoidoscopy; however, they still have several limitations.[2] In addition, the number of people undergoing the screening and repeated tests tend to avoid these tests because of their invasiveness and unpleasantness.[3] Therefore, finding specific and sensitive biomarkers from biological specimen (e.g., blood) would be ideals and relatively a non-invasive alternative way for CRC detection. Until now, carcinoembryogenic antigen (CEA) is the only biomarker approved for clinical use to monitor CRC recurrence. However, it remains a problem because of its limited levels of sensitivity and specificity.[2]

Glycosylation is a common post-translational modification of intra- and extra-cellular proteins, which provides a dynamic mechanism for regulating protein functions and many cellular processes.[4] In recent year, growing evidence reveals that changes in glycosylation are associated with tumor malignancy. Alteration in glycosylation patterns have been observed in human cancers such as prostate cancer,[5] breast cancer,[6] and CRC.[7] Thus, tumor-related glycosylation may serve as a distinct feature of cancer cells and provide novel biomarkers as well as drug targets.[8]

Nowadays, there is an increasing use of lectin affinity chromatography to enrich glycoproteins and enhance the detection of low abundant proteins in biological samples, particularly in serum. As most proteins in serum are glycosylated, the feature of lectins which can bind reversibly to specific glycan structures on glycoproteins can be employed to enrich serum glycoproteins. Wheat germ agglutinin (WGA), a lectin binding protein that has high affinity for GlcNAc/sialic acid residues, has become an exploited tool in many studies for capturing glycoproteins in the biological samples.[9,10]

In this study, we aimed to identify and quantify serum glycoproteins from metastatic CRC patients and healthy controls using WGA affinity chromatography enrichment and label-free protein relative quantitation. WGA-bound glycoproteins identified from two groups were compared by ProteinScape software which based on comparing peak intensity exported from nano-LC-MS scan with t-test.[11] One of glycoproteins displayed the difference in their expression levels between two groups were chosen and validated using immunoblotting and enzyme-linked immunosorbent assays (ELISA).

Methodology

Serum samples

A set of pooled serum samples from 7 CRC patients with metastatic state (age range 50-65 years) and 7 age-matched healthy controls (without hypertension, dyslipidemia, and diabetes) were obtained from Pramongkutklao Hospital (Bangkok, Thailand). The samples were stored at -80˚C until processing. Serum of patients was collected after diagnosis as CRC before any treatment. The informed consent and the research protocol for collecting and handling samples were approved by the Institutional Review Board of the Royal Thai Army Medical Department, Thailand (S012h/56).

WGA glycoprotein Enrichment

Wheat germ agglutinin (WGA) kit (Pierce™ Thermo Scientific) was used to enrich glycoproteins from crude serum. WGA glycoprotein enrichment was performed using 30 µL of crude serum with following the manufacturer's protocol. Protein concentration of crude serum, WGA enriched and unbound fractions were determined using Bradford assay (Bio-Rad).

SDS-PAGE

Equal protein amount of CRC and normal of crude sera, WGA-bound and WGA-unbound serum fractions were separated by 10% SDS-PAGE (TGX Stain-Free™ FastCast™) (Bio-rad) at constant current of 10 mA/gel till the dye front reaching the bottom of the gel. Total proteins were visualized using Gel Doc™ EZ Imager (Bio-rad).

In-solution tryptic digestion

WGA-bound glycoproteins were diluted with 50 mM Ammonium bicarbonate (NH_4HCO_3) to a final concentration of 0.5 mg/mL. The samples (0.5 µg) were reduced with 10 mM dithiothreitol (DTT) at 95˚C for 5 min and alkylated with 20 mM iodoacetamide (IAA) at RT for 30 min in darkness. Then samples were digested with trypsin (Promega) at ratio of trypsin:protein = 1:50 at 37˚C overnight. The trypsinized peptides were desalted using

ZipTip$_{C18}$ (Merck Millipore), dried using SpeedVac, and stored at -20˚C until the day of analysis by LC-MS/MS.

LC-MS/MS analysis

Desalted peptides were analyzed by LC-MS/MS that use a nanoflow liquid chromatography (Bruker) coupled with the ion trap mass spectrophotometry (AmaZon speed, Bruker). The peptides (100 ng) were loaded onto a nanoflow liquid chromatography (75 µm id x 100 mm C18 EASY-nLC™ column (Thermo Scientific)). A 120-min gradient separation at 0.3 nl/min was performed using solution A (0.1% formic acid in water) and solution B (0.1% formic acid in ACN). Each sample was injected three times with the same method, providing three technical replicates per sample. For MS/MS analyses, low energy CID was performed to generate fragment ions. Peak lists of MS/MS spectra obtained from the amaZon speed ion trap were processed using DataAnalysis 4.4 software (Bruker), with the following parameters: compounds (autoMSn) threshold 50000 and maximum number of compounds 20000.

Glycoprotein identification

MS/MS data were analyzed using an in-house Mascot v.2.4.0 search engine (Matrix Science) and SwissProt database (SwissProt_20170118.fasta) and then search results were combined in ProteinScape 3.1 (Bruker) with the following parameters: tryptic peptides with up to 1 missed cleavage sites, peptide tolerance 0.6 Da, fragment tolerance 0.6 Da, 13C=2, instrument type: ESI-TRAP, fixed modifications of cysteine and N-terminal carbamidomethyl and variable modifications of methionine oxidation. All proteins were identified with a false discovery rate (FDR) of 0.05 based on a decoy database search.

Label-free quantitative analysis for WGA-bound glycoprotein in CRC and control sera

Label-free quantification of the digested peptides was performed to compare a relative protein ratio between CRC and control groups. Three MS injections of each group were analyzed using ProfileAnalysis 2.1 (Bruker) for time alignment, integration of chromatographic peak areas for precursor ions, bucket fold changes and t-test calculation. Peptides quantified with ProfileAnalysis 2.1 calculations were exported into ProteinScape 3.1 (Bruker) and mapped to protein IDs based on matching m/z and Rt combinations using deviation threshold = 2 Da for m/z and 2 min for Rt. The peptides that matched the criteria stated above and found in both CRC samples and controls were taken to determine the relative ratio between two groups. Only glycoproteins with differential expression [CRC:normal with ratios >1.5 (up-regulated) and <0.67 (down-regulated)] were considered. The levels of the glycoproteins with ratios between 1.5 and 0.67 were defined as unchanged.

Protein category and data analysis

Glycoprotein identification information was imported to PANTHER database for gene ontology analysis (http//:www.pantherdb.org).

Immunoblot validation of a differently expressed glycoprotein

To validate the protein observation in quantitative MS experiments, immunoblot analysis was performed. Fibronectin was chosen to examine its differential expression level between two groups of serum samples. Protein samples were separated using Stain-free SDS-PAGE and transferred to PVDF membrane. The membrane was blocked with 3% BSA in TBS/T and incubated with a fibronectin antibody (Abcam) at 1:2000 dilution. After washing steps, the membrane was incubated with a HRP conjugated secondary antibody (Dako) at 1:2000 dilution. Protein bands were visualized using Western Bright ECL (Advansta) and the signals were captured by ImageQuant™ LAS 4000 digital imaging system (GE Healthcare).

Measurement of glycoprotein using ELISA

Fibronectin Human Simple Step ELISA kit was purchased from Abcam. Fibronectin of crude sera, WGA-bound and WGA-unbound serum fractions were measured according to the manufacturing instruction with the dilutions at 1:20000, 1:500, and 1:200, respectively. The plates were read with the absorbance at 450 nm using a microplate spectrophotometer (Molecular Devices).

Statistical analysis

The statistical analysis was analyzed using unpaired Student's *t*-test to test for the difference between two groups. The statistical significance was defined as p value <0.05.

Results and Discussion

Enrichment of serum glycoproteins using WGA kit

Pooled samples of 7 metastatic CRC patients and 7 healthy controls were used in this study. Protein samples including crude sera, WGA-bound and WGA-unbound serum fractions were separated by 10% stain-free SDS-PAGE. The result revealed a distinct pattern of proteins observed among three types of sample (Fig. 1). After glycoprotein enrichment, high abundant proteins at around 50-70 kDa presented in crude sera were removed in WGA-bound serum fractions while these proteins were major proteins found in WGA-unbound serum fractions. Therefore, WGA lectin enriched the specific glycoproteins in serum. In addition, the lectin also partially removed the high abundant proteins i.e. albumin, resulting in unmasking of low abundant proteins in serum. However, we could not see the difference of protein bands between sera of metastatic CRC patients and controls.

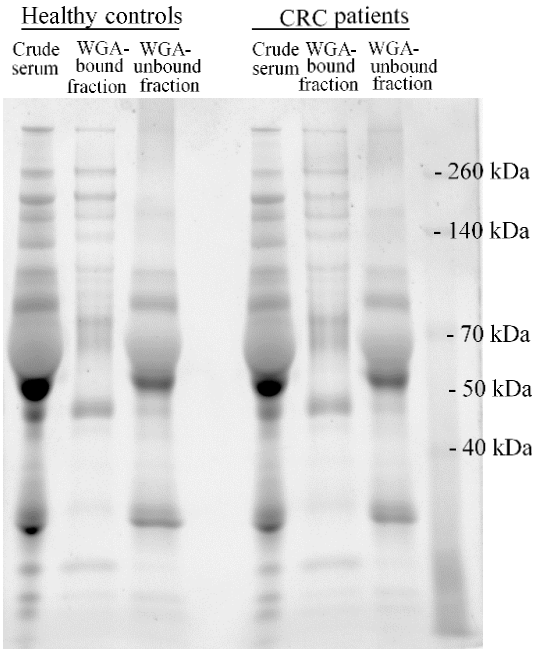


Figure 1. The pattern of total proteins from different fractions of pooled sera of metastatic CRC patients and healthy controls. Crude serum (50 µg), WGA-bound serum fraction (5 µg), and WGA-unbound serum fraction (30 µg) of pooled CRC patients and healthy controls were separated on 10% Stain-free SDS-PAGE. Total proteins were activated and visualized using Gel Doc™ EZ Imager.

Relative glycoprotein quantification between CRC and control sera

Using label-free quantitative proteomics, 37 WGA-bound glycoproteins were identified by combining all data from pooled sera of two groups. Of 37 proteins, only 24 proteins were matched and compared between CRC and control groups. The identified glycoproteins and relative ratios for each glycoprotein expression between CRC and control groups determined by ProteinScape were shown in Table 1. According to 1.5-ratio cutoff threshold, only 3 glycoproteins (complement C4-A, Ig mu chain C region, and hemoglobin subunit beta) were up-regulated whereas 5 glycoproteins (Ig lamda-2 chain C regions, alpha-1-acid glycoprotein 1, haptoglobin, Ig gamma-1 chain C region, and fibronectin) were down-regulated in the CRC serum compared to those in the control, respectively.

Table 1. WGA-bound serum glycoproteins identified from CRC patients and healthy controls using nano-LC-ESI-MS/MS and ProteinScape analysis.

No.	Protein	MW [kDa]	pI	Scores	Relative ratio (CRC:Normal)
1	Complement C4-A	192.7	6.7	924.9	3.56
2	Ig mu chain C region	49.3	6.3	566.9	1.72
3	Hemoglobin subunit beta	16.0	6.7	355.1	1.63
4	Inter-alpha-trypsin inhibitor heavy chain H2	106.4	6.4	146.8	1.48
5	Alpha-1-antichymotrypsin	47.6	5.3	270.2	1.40
6	Complement factor H	139.0	6.2	224.1	1.25
7	Hemoglobin subunit alpha	15.2	8.7	152.8	1.24
8	Serum albumin	69.3	5.9	277.7	1.22
9	Vitronectin	54.3	5.6	62.4	1.21
10	Ig alpha-1 chain C region	37.6	6.1	98	1.17
11	Immunoglobulin lambda-like polypeptide 5	23.0	9.1	114.9	1.17
12	Hemopexin	51.6	6.5	217.5	1.12
13	Inter-alpha-trypsin inhibitor heavy chain H4	103.3	6.5	347.9	1.11
14	Alpha-2-HS-glycoprotein	39.3	5.4	166.3	0.95
15	C4b-binding protein alpha chain	67.0	7.2	234.7	0.93
16	Plasma protease C1 inhibitor	55.1	6.1	123.8	0.89
17	Complement C3	187.0	6.0	1582.3	0.80
18	Alpha-2-macroglobulin	163.2	6.0	1745	0.78
19	Ceruloplasmin	122.1	5.4	447.2	0.71
20	Ig lambda-2 chain C regions	11.3	6.9	87.8	0.66
21	Alpha-1-acid glycoprotein 1	23.5	4.9	110.9	0.64
22	Haptoglobin	45.2	6.1	738.4	0.63
23	Ig gamma-1 chain C region	36.1	8.5	264	0.56
24	Fibronectin	262.5	5.5	260.4	0.52

Gene ontology analysis of glycoproteins

Gene ontology analysis of 24 glycoproteins was performed using PANTHER databases. Cellular component analysis displayed that the majority of these glycoproteins were located in extracellular region (Fig 2A). Another large two portions were belonged to macromolecular complex and membrane proteins. The rest portion was proteins associated to the extracellular matrix. Molecular function analysis showed that the identified proteins were categorized into protein binding and catalytic activities (Fig 2B). Biological process analysis revealed that the proteins were contributed to many biological systems such as cellular process, metabolic process, response to stimulus, biological regulation, and immune system (Fig 2C).

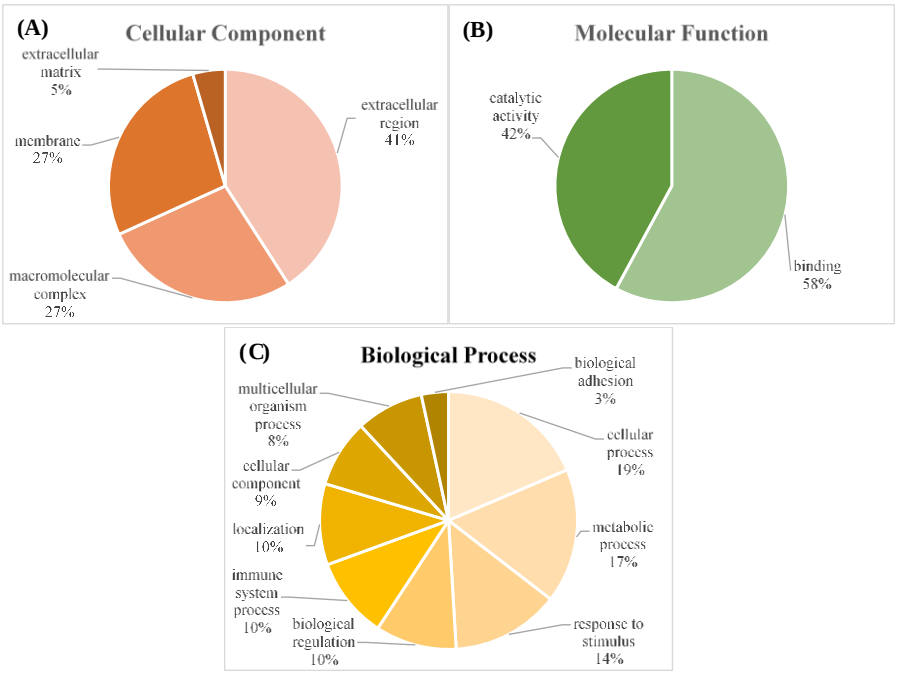


Figure 2. Gene ontology analysis of comparative serum glycoproteins of CRC and control groups. Cellular component (A), molecular function (B), and biological process (C) were predicted by PANTHER.

Validation of fibronectin levels in sera of CRC patients

Among 8 differentially expressed glycoproteins with >1.5 fold changes, complement C4-A showed the highest up-regulation while fibronectin revealed the greatest down-regulation in the CRC serum compared with those in the controls. These glycoproteins were considered to be promising biomarker candidates.

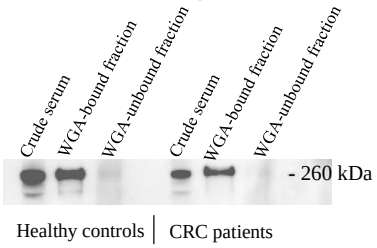
Complement C4-A (C4a) is a cleavage glycoprotein fragment of complement component 4 (C4) that is produced during the activation of the complement cascade. C4a influences in innate immunity.[12] It is a part of anaphylatoxins which participate in several biological responses such as smooth muscle contraction, vasodilation, vascular permeability, chemotaxis, phagocytosis, and inflammation. Complement activation in the tumor microenvironment was reported to enhance tumor growth and increases metastasis.[13]

Fibronectin is a glycoprotein, which presented mainly in the blood, but is also found in the extracellular matrix.[14] It plays important roles in cell adhesion, growth, migration as well as differentiation, which are mediated through integrin signaling. In addition, fibronectin is involved in wound healing and blood coagulation processes.[15,16] It has also been implicated in cancer-associated process by promoting tumor growth, invasion, and metastasis.[15]

Here, in order to verify our label-free quantitative analysis, fibronectin was chosen for validation using immunoblots and ELISA. The results of label-free quantitation indicated that fibronectin was down-regulated in the sera of metastatic CRC patients in comparison to that of control subjects. This was confirmed using immunoblots (Fig 3A). Immunoblot of fibronectin clearly showed that fibronectin level was decreased in crude sera and WGA-bound serum fraction of pooled CRC patients compared to that of pooled normal controls, while it was undetectable in WGA-unbound serum fraction. Interestingly, the immunoblot analysis revealed that the ratio of fibronectin was lower in crude sera (CRC:Control with ratio 0.34) and WGA-bound sera fraction (CRC:Control with ratio 0.41) (Fig 3B). This reduction was consistent with the results obtained from the label-free quantitation (CRC:Control in WGA-bound serum fraction with ratio 0.52).

We also validated the levels of fibronectin in pooled sera of metastasis CRC patients and controls using a commercial ELISA kit. We found that the level of fibronectin in crude serum and WGA-bound serum fraction significantly lower in CRC patients as compared to those of controls (Fig 3C-D). While the level of fibronectin in WGA-unbound serum fraction was not different between CRC and control groups (Fig 3E).

(A) Immunoblot analysis of fibronectin

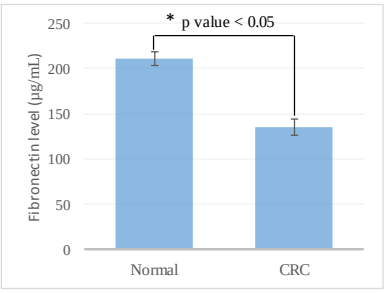


(B) Intensity of fibronectin from immunoblot

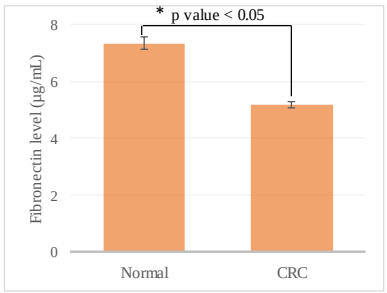
Healthy controls			CRC patients			Ratio (CRC/control)		
Crude sera	WGA-bound fractions	WGA-unbound fractions	Crude sera	WGA-bound fractions	WGA-unbound fractions	Crude sera	WGA-bound fractions	WGA-unbound fractions
1.37×10^8	1.10×10^8	ND	4.62×10^7	4.46×10^7	ND	0.34	0.41	ND

* ND indiacted that intensity in WGA-unbound fractions in CRC and control group would be undetecable (cannot see by naked eyes)

(C) Crude sera



(D) WGA-bound fractions



(E) WGA-unbound fractions

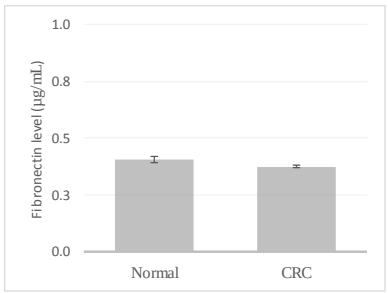


Figure 3. The levels of fibronectin in three different pooled serum fractions of metastatic CRC patients and healthy controls. Representative immunoblot of fibronectin in crude sera, WGA-bound and WGA-unbound fractions (A). Crude serum (50 µg), WGA-bound serum fraction (5 µg), and WGA-unbound serum fraction (30 µg) were used and immunoblot was performed as described in the method. The table showed the levels of fibronectin in each types of sample determined by immunoblot analysis (B). The fibronectin levels measured by ELISA in crude sera (C), WGA-bound fractions (D), and WGA-unbound fractions (E).

Taken together, this study demonstrated the use of WGA affinity chromatography integrated with label-free quantitative proteomics to discover the low-abundance glycoproteins presented in serum which may serve as CRC biomarkers. However, our finding was inconsistent with previous studies in the literature, where elevated fibronectin levels has been observed in plasma[17] and in serum[14] of CRC patients as compared to controls, and also be correlated to cancer progression. It is noted that fibronectin levels in human sera may be variable, probably due to the physiological states (i.e. vascular tissue damage and inflammation) or in some diseases such as atherosclerosis and ischemic heart disease, resulting in elevation of fibronectin levels in the blood circulation.[18]

As CRC is a predominant disease of elderly people, age-matched controls are very important for finding potential biomarkers. However, this study was primarily focused on finding differentially expressed serum glycoproteins between CRC patients with metastatic state and age-matched healthy controls and verifying of fibronectin as a metastatic CRC biomarker. These may limit our ability to find biomarkers for early stages of the disease. Moreover, other differentially expressed serum glycoproteins such as complement C4-A should be further validated. In addition, the proteomic data and validation presented here were from the pooled samples of two groups which may not represent for each individual sample. The further experiments tested in individual sample must be performed in order to observe the variation among sample groups. Also, larger sample size and patient samples with different CRC stages i.e. non-metastatic stage and other diseases as well as a larger set of statistical

analysis are needed to done in order to evaluate whether they have diagnostic capabilities as potential biomarkers.

Conclusion

In this study, wheat germ agglutinin (WGA) enrichment and label-free quantitative proteomic approach were applied to identify differentially expressed serum glycoproteins between CRC patients and age-matched healthy controls. Total 8 glycoproteins were identified to be differentially expressed in sera of patients with CRC. Among those glycoproteins, significantly decreased in fibronectin level in sera of CRC patients was confirmed using immunoblots and ELISA. This may provide an alternative way to find potential CRC biomarkers which could be translated for the clinical practice in the future.

References

1. Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, et al. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. International Journal of Cancer. 2015;136(5):E359-E86.
2. Gonzalez-Pons M, Cruz-Correa M. Colorectal Cancer Biomarkers: Where Are We Now? BioMed Research International. 2015;2015:14.
3. Corbo C, Cevenini A, Salvatore F. Biomarker discovery by proteomics-based approaches for early detection and personalized medicine in colorectal cancer. PROTEOMICS – Clinical Applications. 2017;11(5-6):1600072-n/a.
4. Moremen KW, Tiemeyer M, Nairn AV. Vertebrate protein glycosylation: diversity, synthesis and function. Nature reviews Molecular cell biology. 2012;13(7):448-62.
5. Peracaula R, Tabarés G, Royle L, Harvey DJ, Dwek RA, Rudd PM, et al. Altered glycosylation pattern allows the distinction between prostate-specific antigen (PSA) from normal and tumor origins. Glycobiology. 2003;13(6):457-70.
6. Ng RC, Roberts AN, Wilson RG, Latner AL, Turner GA. Analyses of protein extracts of human breast cancers: changes in glycoprotein content linked to the malignant phenotype. British Journal of Cancer. 1987;55(3):249-54.
7. Kellokumpu S, Sormunen R, Kellokumpu I. Abnormal glycosylation and altered Golgi structure in colorectal cancer: dependence on intra-Golgi pH. FEBS Letters. 2002;516(1-3):217-24.
8. Stowell SR, Ju T, Cummings RD. Protein Glycosylation in Cancer. Annual review of pathology. 2015;10:473-510.
9. Lee LY, Hincapie M, Packer N, Baker MS, Hancock WS, Fanayan S. An optimized approach for enrichment of glycoproteins from cell culture lysates using native multi-lectin affinity chromatography. Journal of Separation Science. 2012;35(18):2445-52.
10. Schwefel D, Maierhofer C, Beck JG, Seeberger S, Diederichs K, Möller HM, et al. Structural Basis of Multivalent Binding to Wheat Germ Agglutinin. Journal of the American Chemical Society. 2010;132(25):8704-19.
11. Chang C-T, Yang C-Y, Tsai F-J, Lin S-Y, Chen C-J. Mass Spectrometry-Based Proteomic Study Makes High-Density Lipoprotein a Biomarker for Atherosclerotic Vascular Disease. BioMed Research International. 2015;2015:13.
12. Dunkelberger JR, Song W-C. Complement and its role in innate and adaptive immune responses. Cell Research. 2009;20:34.
13. Afshar-Kharghan V. The role of the complement system in cancer. The Journal of Clinical Investigation. 2017;127(3):780-9.
14. Saito N, Nishimura H, Kameoka S. Clinical significance of fibronectin expression in colorectal cancer. Molecular medicine reports. 2008;1(1):77-81.
15. Wang JP, Hielscher A. Fibronectin: How Its Aberrant Expression in Tumors May Improve Therapeutic Targeting. Journal of Cancer. 2017;8(4):674-82.

16. Yi W, Xiao E, Ding R, Luo P, Yang Y. High expression of fibronectin is associated with poor prognosis, cell proliferation and malignancy via the NF-κB/p53-apoptosis signaling pathway in colorectal cancer. Oncology Reports. 2016;36(6):3145-53.
17. Choate JJ, Mosher DF. Fibronectin concentration in plasma of patients with breast cancer, colon cancer, and acute leukemia. Cancer. 1983;51(6):1142-7.
18. To WS, Midwood KS. Plasma and cellular fibronectin: distinct and independent functions during tissue repair. Fibrogenesis & Tissue Repair. 2011;4(1):21.