

Glycoproteomic Analysis of Human Hepatoblastoma Cell Lines Using Glycopeptide Capture and Mass Spectrometry

Qinze Wang†, Fei Zhao†, Xiuxuan Sun, Shisheng Sun, Hanjie Yu, Yaogang Zhong, Yuan Hu, Zhuo Chen, Peixin Zhang and Zheng Li*

Laboratory for Functional Glycomics, College of Life Sciences, Northwest University, Xi'an 710069, China

†These authors contributed equally to this work

Abstract

Hepatoblastoma (HB) is the most common form of liver tumour in infants and children. The cancer cell line HepG2 and normal hepatocellular cell line L02 are valuable cell models and are already widely used in the world while their glycopeptide profiles are still unknown. This study focuses on N-linked glycoproteomic analysis of these two cell lines using mass spectrometry. Using two complementary approaches, Hydrazide reaction and hydrophilic affinity solid phase extraction methods, almost 400 glycosylation sites were identified from the two cell lines. Functional annotation suggests that N-glycoproteins with molecular binding ability were enriched in HepG2 cells compared to normal liver cells.

Keywords: Hepatoblastoma; Glycoproteome; Mass spectrometry; Hydrazide reaction; Hydrophilic affinity

Abbreviations

HSC: Hepatic Stellate Cell; HB: Hepatoblastoma; HCC: Hepatocellular Carcinoma; IGF2: Insulin-Like Growth Factor 2; DLK1: Protein Delta Homolog 1; TGF- β 1: Transforming Growth Factor- β 1; MALAT1: Metastasis-Associated Lung Adenocarcinoma Transcript 1; MIG6: ERBB Receptor Feedback Inhibitor 1; HEPES: 4-(2-Hydroxyethyl)-1-Piperazine \bar{e} thanesulfonic Acid

Introduction

Hepatoblastoma (HB) is the most common form of liver tumor in infants and children [1] with the age of 2 months to 3 years old [2]. Compared to Hepatocellular carcinoma (HCC), which majorly affects adults and is caused by the infection of hepatitis virus. Hepatoblastoma is a result of accumulating mutations during the growth of liver cells. Although several genetic conditions are associated with the increased risk of developing hepatoblastoma, such as Beckwith-Wiedemann syndrome [3] and Familial Adenomatous Polyposis [4], the pathological details of hepatoblastoma are yet to be determined. A transcriptomic and genomic analysis demonstrated that the gene expression patterns between HB and HCC are different [5]. Several genes are over expressed in HB compared to in HCC, such as IGF2, Fibronectin, DLK1, TGF- β 1, MALAT1 and MIG6.

HepG2 cell line is derived from hepatoblastoma and has been widely used as liver cancer model. For example, Blazquez et al. [6] transfected HepG2 cells with CCAAT; enhancer-binding protein β (CEBP β), hepatocyte nuclear factor 4 α (HNF4 α), and constitutive androstane receptor (CAR) to create a stable model to study bile acid biosynthesis. Investigations at proteome level using HepG2 cells are becoming popular as the progress of proteomics. Pattanakitsakul et al. [7] analyzed the host responses in HepG2 cells during Dengue virus infection and 17 differentially expressed proteins were identified by 2-D PAGE. Tong et al. [8] identified Hepatitis B virus X protein as an inducer for hyper methylation in HepG2 cells.

Glycosylation has long been recognized as one of the most common post-translational modifications of proteins [9]. Glycosylation not only plays a key role in protein stability, solubility, folding, and activity, but also participates in cellular processes such as protein trafficking, sorting and molecular recognition [10-12]. Aberrant glycosylation is a fundamental characteristic of cancer and other diseases [12-14], and many clinical biomarkers and therapeutic targets are glycoproteins [15,16]. N-linked glycoproteins are prevalent on the extracellular

side of the plasma membrane, in secreted proteins, and in proteins retained in body fluids [17]. These membrane glycoproteins are easily accessible to therapeutic drugs, antibodies, and ligands; and the secreted glycoproteins in the body fluids are useful in patient diagnosis or clinical management. Therefore, glycoproteomic analysis of HepG2 and L02 cells will not only provide a deeper understanding of hepatoblastoma but also may lead to potential cancer biomarkers.

Many glycoprotein and glycopeptide enrichment methods have been developed for mass spectrometric analysis. These methods include lectin affinity [18,19], immunoaffinity [20,21], boric acid chemistry [22,23], size exclusion [24,25], hydrophilic interaction [26-28], and hydrazide chemistry [29,30]. Among these, hydrophilic interaction and hydrazide chemistry methods are two important methods that can theoretically isolate all types of N-glycopeptides regardless of their attached glycans. Hydrophilic interaction solid-phase extraction method [31], which takes advantage of the higher hydrophilicity of glycopeptides vs. non-glycopeptide, is able to efficiently enrich glycopeptide from peptide mixtures. Hydrazide chemistry isolates glycopeptides from complex samples through the formation of hydrazone bonds between the hydrazide functional groups on solid-phase support and aldehydes from oxidized glycans. Compared to the methods mentioned above to extract all N-glycopeptides, selective extraction of the sub-glycoproteome with glycans containing terminal sialic acid can be performed by a modified hydrazide chemistry method [32], which utilizes a reduced concentration of NaIO₄ and lowered reaction temperature at oxidation step. It has been reported that these methods were complementary and using a combination of these method could lead to a more comprehensive glycoproteome profiling than using individual method [33].

In this study, we used a combination of hydrazide chemistry and hydrophilic interaction methods to extract N-glycopeptides from

*Corresponding author: Zheng Li, Laboratory for Functional Glycomics, College of Life Sciences, Northwest University, 229 Taibai Beilu, Xi'an 710069, P. R. China, Tel: +86-29-88302411; E-mail: zhengli@nwu.edu.cn

Received: June 11, 2015; Accepted: July 03, 2015; Published: July 15, 2015

Citation: Wang Q, Zhao F, Sun X, Sun S, Yu H, et al. (2015) Glycoproteomic Analysis of Human Hepatoblastoma Cell Lines Using Glycopeptide Capture and Mass Spectrometry. J Glycomics Lipidomics 5: 132. doi:10.4172/2153-0637.1000132

Copyright: © 2015 Wang Q, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

HepG2 cells and L02 cells (Figure 1). The extracted N-glycopeptides were analyzed by LC-MS; MS and N-glycoproteins with the corresponding N-linked glycosylation sites were identified. A series of bioinformatics analysis were performed to annotate the identified glycoproteins. The dataset introduced in this study provides a basis for further functional studies such as analyzing glycan mediated recognition events.

Materials and Methods

Materials

Gibco® RPMI 1640 medium was purchased from Invitrogen (Carlsbad, CA, USA). Non-essential amino acids were from Hyclone (Losan, UT, USA). Fetal bovine serum was from Zhejiang Tianhang Biological Technology Co. Ltd. (Zhejiang, China). Calf serum was from Minhai Biotechnology Co. Ltd. (Beijing, China). 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) was from Qingdao MDBio Biotech Co. Ltd. (Qingdao, China). Affi-Gel® Hz Hydrazide Gel was from Bio-Rad Laboratories (Hercules, CA, USA). Dithiothreitol (DTT) was from GE Healthcare (Waukesha, WI, USA). Sequencing grade modified Trypsin was from Promega (Madison, WI, USA). PNGase F was from New England Biolabs Inc. (Beverly, MA, USA). Other compounds such as iodoacetamide (IAA) and urea were from Sigma-Aldrich (St Louis, MO USA). Sep-Pak C18 1 cc Vac Cartridge, 50 mg Sorbent per Cartridge was from Waters (Milford, MA, USA).

Cell culture and sample preparation

Cell cultures of Human hepatoblastoma cell line HepG2 and human normal liver cell L02 were grown with RPMI 1640 medium supplemented with 10% fetal bovine serum and antibiotics. Cells were harvested at 90% confluence. To remove the possible contaminating proteins from culture medium, cells were washed with cold 0.1 M phosphate buffered saline (PBS), pH 7.4 for 3 times. Cells were then lysed with a buffer consisting of 8 M urea, 0.1 M sodium bicarbonate, and 0.1% sodium dodecyl sulfate (SDS) on ice for 30 minutes. Lysates were sonicated using an ultrasonic cell disruptor (Xinzhi, NingBo, China) on 100W for 5 minutes (60% duty cycle). The sample was clarified by centrifugation at 10,000 × g for 60 min. The total amount of proteins from supernatant was quantified by BCA method and kept at -80°C for future use.

Protein digestion

0.5 mg/ml total proteins were reduced by 5 mM DTT at 37°C for 60 min and alkylated by 20 mM iodoacetamide at room temperature in the dark for 60 min. After the urea concentration was diluted below 2 M with a 0.1M NH₄HCO₃ solution, 10 µg sequencing grade modified trypsin (Promega, USA) was introduced and the glycoproteins were incubated at 37°C overnight. The resulting peptides were purified by Sep-Pak® Vac C18 cartridge according to manufacturer's instruction. Peptides were dried with a Freeze Dryer (Christ Alpha 1-4 LSC, Germany) and stored at -20°C for future use.

Solid phase extraction of total N-glycopeptide (SPEG)

Total N-glycopeptides of plasma proteins were enriched according to Tian et al. [30]. Briefly, the *cis*-diols on glycans of tryptic glycopeptides were oxidized with 10 mM NaIO₄ in 0.1% trifluoroacetic acid (TFA) at room temperature in dark for 60 minutes. Then, excess amount of NaIO₄ were removed by Sep-Pak® Vac C18 cartridge and the oxidized glycopeptides were mixed with 25 µl hydrazide resin. After an overnight reaction, the oxidized glycopeptides were covalently attached to the hydrazide resin. N-glycopeptides were released by a glycan cleaving reaction using 1000 U of Peptide-N-Glycosidase F in 25mM NH₄HCO₃ at pH 7.8 overnight following extensively washing hydrazide resin-glycopeptides complexes to remove non-glycopeptides. The formerly glycopeptides were then purified, dried and stored at -20°C for LC-MS; MS analysis.

Enrichment of N-glycopeptide by hydrophilic affinity solid phase extraction method

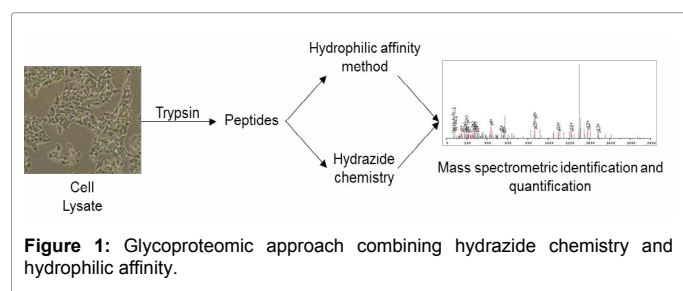
Hydrophilic affinity solid phase extraction method took advantage of the difference in hydrophilicity between normal peptides and glycopeptides to enrich glycopeptides. 20 µl of Sepharose CL-4B resin was cleaned using Washing Buffer (n-Butanol: ethanol: water = 5: 1:1 (v:v:v)). 0.5 mg tryptic glycopeptides were dissolved in 500 µl Loading Buffer (n-Butanol: ethanol: water = 5: 1:1 (v:v:v) with 1mM MnCl₂) and mixed with Sepharose CL-4B resin. The mixture was shocked at room temperature for 45 minutes. Glycopeptides were released by Elution buffer (ethanol: water = 1:1 (v:v)) following washing the resin twice with Washing buffer to remove non-glycopeptides. Finally, the enriched glycopeptides were subjected to PNGase F digestion and kept at -20°C for LC-MS; MS analysis.

LC-MS; MS analysis

The previously isolated glycopeptides were separated on Easy-nLC II HPLC system (ThermoFisher, USA) with a C18 EASY-Column (10cm, ID 75 µm, 3 µm, C18) and identified using LTQ Orbitrap XL mass spectrometer (ThermoFisher, USA). The nano-LC separation was performed at a flow rate of 300 nl/min. The eluents used for the LC were (A) 99.9% H₂O; 0.1% FA and (B) 99.9% ACN; 0.1% FA. A gradient was from 3% B to 9% B for 10 min, from 10% B to 41% B for 70 min and from 41% B to 81% B for 15 min, followed by holding at 81% B for 25 min. Collision induced dissociation (CID) was used to generate daughter ions from 10 parent ions of each MS spectrum, and dynamic exclusion was enabled. Spray voltage was set at 1.95 kV and the capillary temperature was 200°C.

Database searching and identification of glycosylation sites

MS;MS spectra were searched against human IPI database using Sequest algorithm of Protein Discoverer (version 1.2., Thermo Fisher, USA) with the following parameters. Precursor Mass Tolerance was set as 10 ppm and Fragment Mass Tolerance was 50 mmu. Trypsin was used as protease and two missed cleavage sites were allowed. Carbamido methylation (+57.021 Da (C)) was set as fixed modification. Oxidation (+15.995 Da (M)) and Deamidation (+0.984 Da (N, Q)) were set as dynamic modifications. A confidence level of 99% was used to generate the identification list. The N-glycosylation sites were confirmed by a deamidation event at asparagine residue in the typical glycosylation motif, which is N-X-T; S; C-X (X can be any amino acid residue except proline). To further analyze the frequency of amino acids at #2 and #4 positions of N-glycosylation, a motif analysis was performed using WebLogo [34]. The frequency of amino acid showing at each position was analyzed manually.



Label free relative quantification of glycopeptides

Expression level of identified glycoproteins between HepG2 and L02 cells were estimated using spectral counting methods as described by Tian et al. [35]. The quantification was performed at protein expression level. Total numbers of spectra for each identified glycoprotein was summed. Glycoproteins with spectral counts of at least 2 were used for quantification. Ratios of total numbers of spectra of same glycoprotein from HepG2 vs. L02 were reported as the relative expression level. For the glycoproteins that were only identified in HepG2 cells, the ratios were arbitrarily marked as 100.

Gene ontology analysis of identified glycoproteins

Gene Ontology (GO) analysis was performed according to the standard procedure of Blast2GO [36] to gain insight about identified glycoproteins in three aspects, namely cellular distribution, molecular function, and potential biological processes that glycoproteins participated in. The differences in GO terms between of the glycoproteins from HepG2 cell line and L01 healthy liver cells were identified.

Results

Glycopeptides and glycoproteins identified from HepG2 and L02 cells

In this research, glycoproteins were released from cell lysate and digested into tryptic peptide mixtures. The enrichment of glycopeptides was performed using hydrazide chemistry and hydrophilic affinity methods. 394 non-redundant *N*-glycopeptides with 398 glycosylation sites from 237 *N*-glycoproteins were identified from HepG2 and L02 cells (Supplementary Table S1). Of these, 256 *N*-glycopeptides, from 164 *N*-glycoproteins, were identified from HepG2 cells. 331 *N*-glycopeptides, from 200 *N*-glycoproteins, were identified from L02

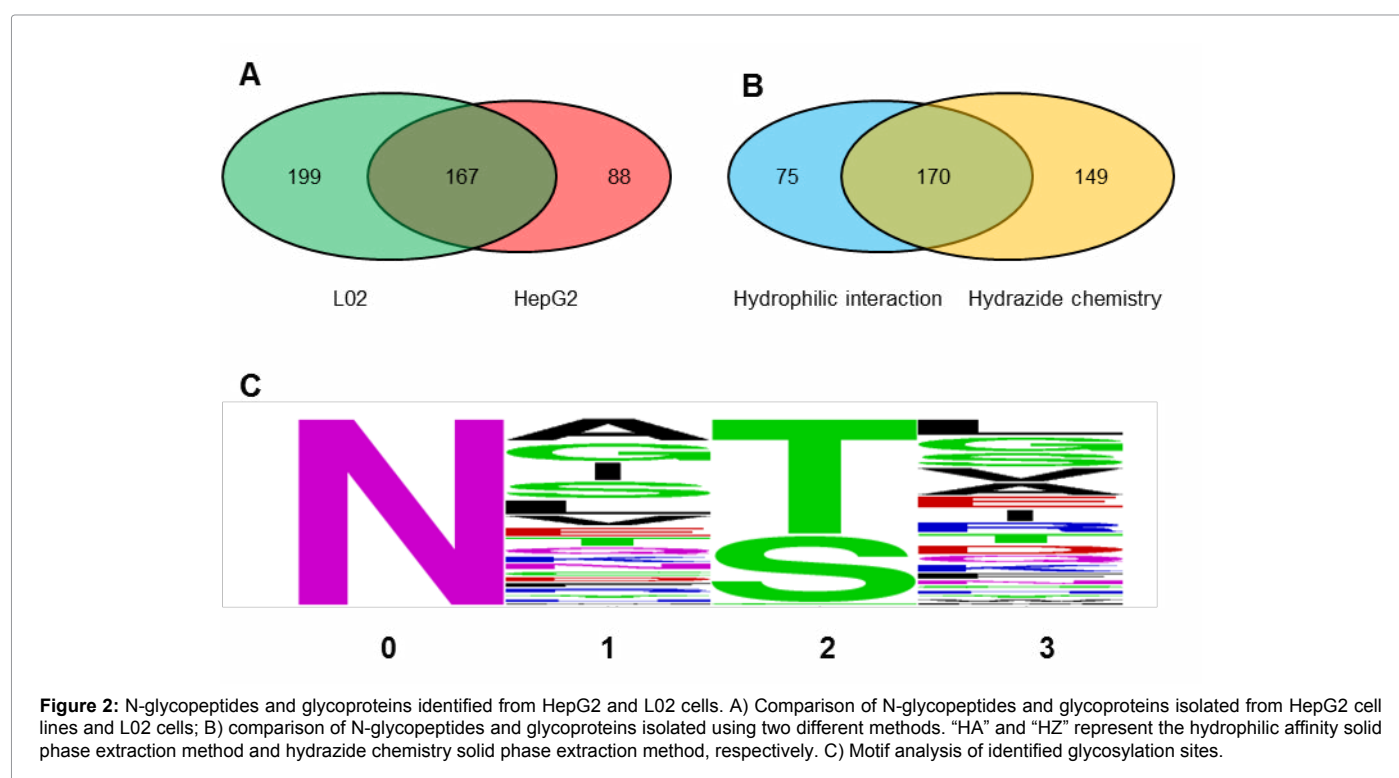
cells. 191 glycopeptides were confirmed in both cell types while 65 glycoproteins were shown in HepG2 cells only (Figure 2A).

Hydrazide chemistry and hydrophilic affinity are complementary approaches to enrich glycopeptides. In the current research, 170 glycopeptides were identified by both methods, 75 glycopeptides were identified by hydrophilic affinity method only, and 149 glycopeptides were identified by hydrazide chemistry method only (Figure 2B).

To confirm the known *N*-glycosylation sites and to identify new *N*-glycosylation sites, each glycosylation site was mapped to the corresponding protein in Uniprot database [37]. 188 *N*-glycosylation sites (47.2%) were newly identified in the current research and 210 known glycosylation sites (52.8%) were confirmed.

Dual identity of glycoproteins; glycan binding proteins

Interestingly, 11 *N*-linked Glycoproteins identified from HepG2 cell line and/or L02 cell line have carbohydrate binding ability according to the annotation from Uniprot database [37]. These glycoproteins include Lysosomal alpha-mannosidase (O00754), Attractin (O75882), Thrombospondin-1 (P07996), Cation-independent mannose-6-phosphate receptor (P11717), Cation-dependent mannose-6-phosphate receptor (P20645), CD44 antigen (P16070), Neural cell adhesion molecule L1 (CD171 antigen, P32004), Basigin (CD147 P35613), Follistatin-related protein 1 (Q12841), Nodal modulator 1 (Q15155), C-type mannose receptor 2 (CD280 Q9UBG0) (Table 1). The dual identity of those glycoprotein; glycobinding proteins may suggest the importance of a glycosylation based regulation network. For example, CD44 antigen, located on the surface of a cell, is the receptor for hyaluronic acid and regulates the cell-cell, cell-matrix interaction, which is crucial for cell migration and tumor growth. C-type mannose receptor 2 has calcium dependent glycol-binding ability and it participates in endocytosis of glycosylated ligand via clathrin [33].



No.	Uniprot Accession	Name	Identified Glycopeptides	HepG2 HZ	HepG2 HA	L02 HZ	L02 HA
1	O00754	Lysosomal alpha-mannosidase	AnLTWSVK			2	
2	O00754	Lysosomal alpha-mannosidase	LnQTEPVAGNYYPVNTR		6	7	
3	O75882	Attractin	IDSTGnVTNELR		1	4	
4	P07996	Thrombospondin-1	VVnSTTGPGHEHLR	4	2	4	4
5	P11717	Cation-independent mannose-6-phosphate receptor	MnFTGGDTcHK		2	2	1
6	P11717	Cation-independent mannose-6-phosphate receptor	mSVINFEcnK-TA			6	
7	P11717	Cation-independent mannose-6-phosphate receptor	nGSSIVDLSPLIHR		2		2
8	P11717	Cation-independent mannose-6-phosphate receptor	TnITLVcKPGDLESAPVLR			2	1
9	P16070	CD44 antigen	AFnSTLPTMAQMEK	2	9	6	8
10	P20645	Cation-dependent mannose-6-phosphate receptor	EAGNHTSGAGLVQlnK-SN	6	7	7	6
11	P20645	Cation-dependent mannose-6-phosphate receptor	LnETHIFnGSNWIMLIYK				6
12	P32004	Neural cell adhesion molecule L1	DLQAnDTGR	2		3	
13	P32004	Neural cell adhesion molecule L1	FFPYAnGTLGIR			2	
14	P32004	Neural cell adhesion molecule L1	GYnVTYWR	1		2	2
15	P32004	Neural cell adhesion molecule L1	THnLTDLSPHLR		2	2	2
16	P32004	Neural cell adhesion molecule L1	VPGnQTSTTLK	1	3	6	4
17	P35613	Basigin	ALMnGSESR	2	2	1	2
18	P35613	Basigin	ILLTcSLnDSATEVTGHR		6	4	5
19	P35613	Basigin	ITDSEDKALMnGSESR	5	6	8	4
20	Q12841	Follistatin-related protein 1	GSnYSEILDK	2		2	
21	Q15155	Nodal modulator 1	LEnITGTGTYIHAQK	3		5	
22	Q15155	Nodal modulator 1	ENVGIYnLSK	2	1		
23	Q9UBG0	C-type mannose receptor 2	WnDSPcnQSLPSIcK		4		4
24	Q9UBG0	C-type mannose receptor 2	KKPnATAEPTPPDR	2	2		
25	Q9UBG0	C-type mannose receptor 2	KPnATAEPTPPDR	2			
26	Q9UBG0	C-type mannose receptor 2	TSnISKPGTLER	5	4	4	4
27	Q9UBG0	C-type mannose receptor 2	VTPAcnTSLPAQR	2	2	3	2

"n" represents the identified N-linked glycosylation sites; "-" represents the tryptic cleavage site. b) Numbers of corresponding spectra for each peptide with each method. HZ- hydrazide chemistry; HA- hydrophilic affinity

Table 1: Identified N-linked glycoproteins with carbohydrate binding ability.

Certain enzymes of O-glycan biosynthesis are N-glycosylated

Glycosyl transferases are responsible for protein glycosylation. The aberrant glycosylation is due to the altered activity of glycosyl transferases. In this study, four enzymes, glucosyl transferase 1 (Q4G148), UDP glucuronosyl transferase 1 family A1 (P22309), glycotransferase 25 domain containing 1(Q8NBJ5) and procollagen-lysine, 2-oxoglutarate 5-dioxygenase 3 (O60568), in O-glycan biosynthesis were also identified by this research. Glucosyl transferase 1 acts on O-glycosylated Notch epidermal growth factor repeats, thus participating in Notch signaling[38]. The N-glycosylation sites of glucosyl transferase 1 were also identified recently [39]. The ASN347 of UDP glucuronosyl transferase 1 family A1 was confirmed by this research as an N-glycosylation site, which exclusively identified from normal liver cell (L02) but not in cancer cell line (HepG2). UDP glucuronosyl transferase 1 family A1 was known as a major component in toxic compounds metabolism by glucuronidation. Although mutations of UDP glucuronosyl transferase has been associated with Gilbert and Crigler-Najjar type II

syndromes [40], the effect of N-glycosylation on enzyme activity has yet to be determined. O-linked glycosylation, such as O-linked N-acetyl glucosamine, has long been recognized as a widespread signaling mechanism in nucleus and cytoplasm [41,42]. It would be interesting to postulate the existence of a cross-talking mechanism between N-glycosylation and O-glycosylation through modulating the activity and; or substrate specificity of glycotransferases.

Motif analysis for N-glycosylation sites

N-glycosylation is protein sequence dependent. It has been reported that in eukaryotic species, the amino acid at second [43] and fourth [44] positions in Asn-Xxx-Thr; Ser; Cys-Xxx (where Xxx is any amino acid except proline) affect the frequency of glycosylation. The hydroxyl [45] or thiol at the third position is hypothesized to increase the nucleophilicity of the amide group of the asparagine, facilitating the glycosylation. From the motif analysis, neutral amino acids, such as AGILV, are more popular at the second and fourth position than the amino acids with phenol ring, such as WF (Figure 2C).

Relative quantification of identified glycopeptides between HepG2 and L02 cells using spectral counting

The glycopeptide extraction step in this study prevents an absolute quantification of certain protein in each sample due to the following reasons. First, non-glycoproteins are unable to be extracted by solid phase support. This alters the composition of peptide mixture, lowering the protein coverage. Second, formerly glycopeptides from different proteins may have different ionization property. Third, it is common to see miss cleavage during tryptic digestion. Same glycosylation site may exist in different peptides, which have different ionization property in ion source. To overcome these difficulties, label-free quantification has been achieved by non-glycopeptides from glycoproteins [46]. Label-free relative quantification of glycopeptide between samples is plausible using isolated glycopeptides although an absolute quantification is very difficult. For example, spectral counts of identified glycosylation sites have been used to compare the expression level of sialylated glycoproteins in breast cancer [35].

Of 222 glycoproteins quantified in this study, 36 *N*-glycoproteins were uniquely identified from HepG2 while 72 *N*-glycoproteins were uniquely identified from L02 cells. 22 glycoproteins were up regulated (ratio > 2) in HepG2 cells while 32 glycopeptides were down regulated (ratio < 0.5) compared to those in L02 cells. A truncated list and the complete list of quantified glycoproteins were shown in Table 2 and Supplementary Table S2, respectively.

Gene ontology of identified glycoproteins

To understand the biological functions of identified glycoprotein, the identified glycoproteins were mapped into 31,798 gene ontology terms. 218 out of 235 *N*-glycoproteins (92.8%) were annotated. 62 *N*-glycoproteins have catalytic activity. 124 *N*-glycoproteins were located on to organelles and 59 were at extracellular region. More than 70% of the identified *N*-glycoproteins were binding partners. Other enriched molecular function terms include catalytic activity (35.7%), molecular transducer activity (19.7%), and transporter activity (11.0%). In terms of the biological processes, most of the identified *N*-glycoproteins participated in cellular process (61.0%), metabolic process (49.5%) and biological regulation (45.8%) (Figure 3A).

To determine if there was any difference in terms of cellular distribution and biological functions between the identified *N*-glycoproteins from HepG2 and L02 cells, we performed a Gene Ontology Enrichment analysis, which demonstrated 12 significantly different GO terms (with $p < 0.05$ by Fisher's exact test). There was no significant difference of the cellular distribution of *N*-glycoproteins between HepG2 and L02 cells. But an enhanced binding ability in *N*-glycoproteins from HepG2 cells, especially carbohydrate binding ability was observed. Also the receptor activity was elevated (Figure 3B). The regulation of biological quality and cellular process terms were also enriched in glycoproteins from HepG2 cells compared to the glycoproteins from L02 cells (Figure 3C).

Discussion

Liver cells retain the ability of fast proliferation and differentiation during our entire life. That ability also comes with higher risk to accumulate genetic errors, which may lead to uncontrolled proliferation (cancer), in cells compared to other organs. Compared to hepatocellular carcinoma, which majorly caused by viruses, hepatoblastoma provides a perfect model to analyze the cancer pathology and protein interaction networks in cancer cells. Also, the early symptoms of hepatoblastoma in infants may due to unknown developmental errors which need to

be addressed. It has been suspected that carcinogenesis happens even before birth [2].

Post-translational modifications are able to significantly modify the function of proteins. Among various modifications, glycosylation is interesting due to the diversity of the structures of glycan and the drastic difference in molecular properties between peptides and carbohydrates. Carbohydrate turns a biomolecule into an information storage unit and an information processing unit when molecular recognition event happens. In terms of cancer pathology, aberrant glycosylation interrupts the normal information flow within cells or between cells by utilizing unusual glycosylation sites and/or generating abnormal glycan structures. These features may lead to the development of clinical biomarkers and therapeutic targets [15,16]. For example, Golgi membrane protein 1 (Golgi protein 73, GP73) has been proposed as a biomarker for hepatocellular carcinoma with 3 potential glycosylation sites. Glycosylation site N109 was confirmed from HepG2 and L02 cells in this study.

With the rapid progress of targeted glycoproteomics, various enrichment methods for glycoprotein; glycopeptides have been developed. Hydrazide chemistry is the most specific method due to the formation of covalent linkages between glycan and solid phase support during extraction of glycopeptides. Other merits include ease to handle and high reproducibility. It has been widely used in serum glycoproteome research, glycosylation analysis in influenza virus and disease-related glyco-biomarker discovery. Several variants of hydrazide chemistry method exist. For example, extraction of glycoproteins before tryptic digestion or extractions of glycopeptides after tryptic digestion are two commonly used methods. The latter method usually allows for more identification due to reduced steric hindrance effect around the *N*-glycosylation sites.

Label-free quantification methods are gaining popularity in glycoproteomics, although the best quantification methods nowadays remains the ones using stable isotope labeling, due to their convenience and enough accuracy as needed [47]. Spectral counting was first introduced by Liu et al. [48] and remains the most popular method. The accuracy of quantification is directly correlated with spectral counts. In this study, more than 51% of identified glycopeptides had at least 4 spectra matched to them, which provided a relatively good estimation for expression level. To further improve the peptide sampling during mass spectrometry, a label-free quantification of glycoproteins by non-glycopeptides using spectral count has been recently described by Chen et al. [46], with a good correlation between spectral count and glycoprotein content.

The glycoproteomic analysis of human hepatoblastoma cell lines by glycopeptide capture and mass spectrometry described above leads to almost 400 glycosylation sites identified from the two cell lines and more than 1; 3 of them were newly identified thus not included in UniProt database. Several glycoproteins highlighted by semi-quantitation assay may lead not only to new functional studies of individual proteins but also the dynamic network view of post-translational glycosylation. For example, Inactive tyrosine-protein kinase 7, whose glycopeptides were only identified from HepG2 cells, is a protein involved in Wnt signal pathway. It participates in cell adhesion, migration and proliferation and could be a proteolytic target for cancer cell invasion [49]. Nicastrin is an integral membrane protein of gamma-secretase complex in Notch signalling pathway. It is overexpressed in breast cancer and is able to promote invasion [50]. However, 5 glycopeptides, 35 spectra total, of Nicastrin were mapped only in L02 lysate in this study, suggesting a different role of Nicastrin in breast cancer and in hepatoblastoma.

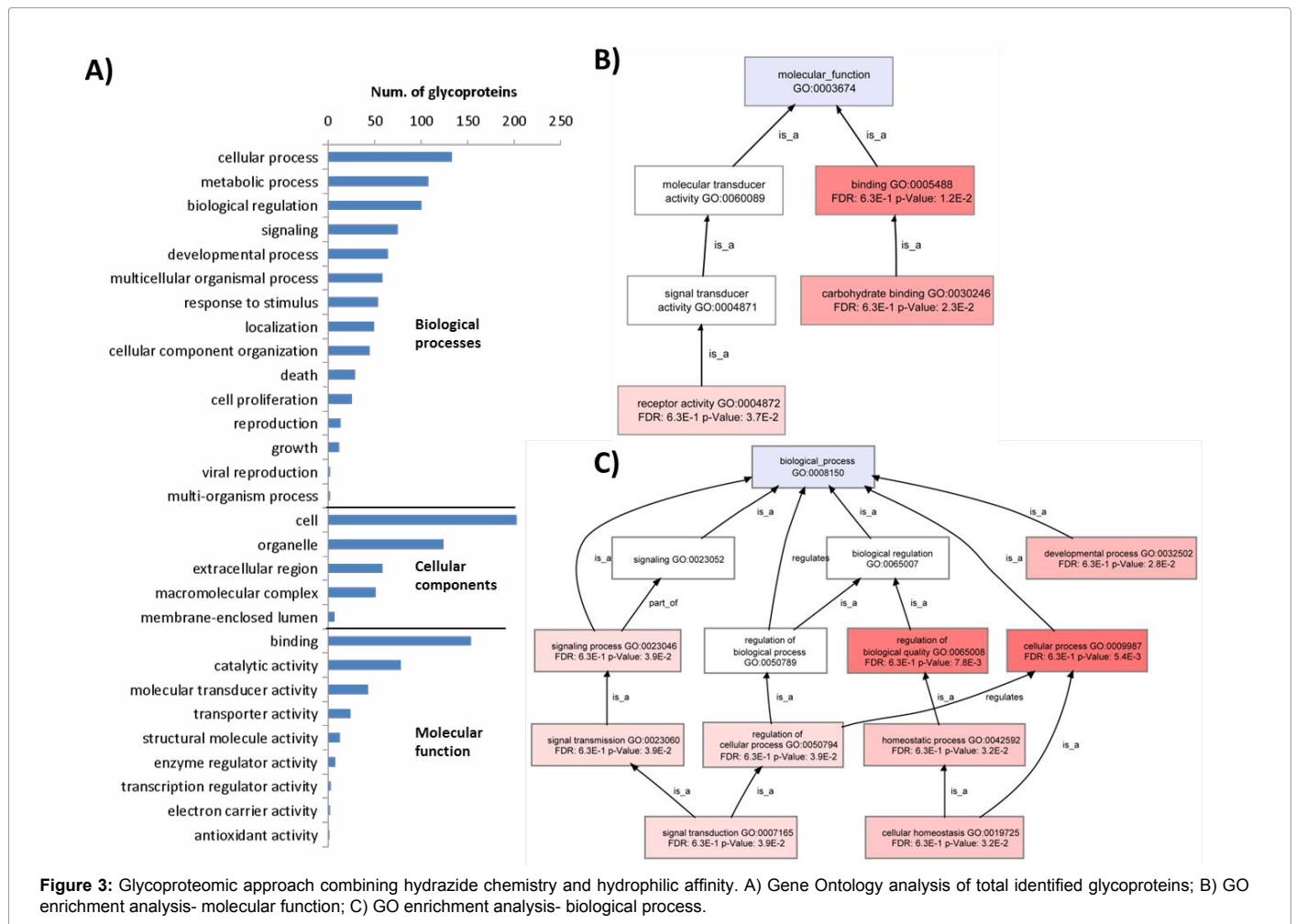


Figure 3: Glycoproteomic approach combining hydrazide chemistry and hydrophilic affinity. A) Gene Ontology analysis of total identified glycoproteins; B) GO enrichment analysis- molecular function; C) GO enrichment analysis- biological process.

No.	Uniprot Accession	Protein Name	Peptide Count	HepG2 Total Spectra	L02 Total Spectra	HepG2 Total Spectra/ L02 Total Spectra
1	Q13308	Inactive tyrosine-protein kinase 7	2	9	0	100
2	P54687	Branched-chain-amino-acid aminotransferase, cytosolic	2	4	0	100
3	Q15818	Neuronal pentraxin-1	1	9	0	100
4	O43399	Tumor protein D54	1	6	0	100
5	Q13433	Zinc transporter ZIP6	1	6	0	100
6	P05186	Alkaline phosphatase, tissue-nonspecific isozyme	1	5	0	100
7	P36941	Tumor necrosis factor receptor superfamily member 3	1	5	0	100
8	O15118	Niemann-Pick C1 protein	1	4	0	100
9	P09923	Intestinal-type alkaline phosphatase	1	4	0	100
10	P17936	Insulin-like growth factor-binding protein 3	1	4	0	100
11	P28300	Protein-lysine 6-oxidase	1	4	0	100
12	Q17RY6	Lymphocyte antigen 6K	1	4	0	100
13	O75197	Low-density lipoprotein receptor-related protein 5	1	3	0	100
14	O75503	cDNA FLJ90628 fis, clone PLACE1003407, highly similar to Ceroid-lipofuscinosis neuronal protein 5	1	3	0	100
15	P01033	Metalloproteinase inhibitor 1	1	3	0	100
16	P28799	Granulins	1	3	0	100
17	P32969	60S ribosomal protein L9	1	3	0	100
18	P43121	Cell surface glycoprotein MUC18	1	3	0	100
19	P50897	Palmitoyl-protein thioesterase 1	1	3	0	100
20	Q00839	Heterogeneous nuclear ribonucleoprotein U	1	3	0	100
21	Q29983	MHC class I polypeptide-related sequence A	1	3	0	100
22	Q99808	Equilibrative nucleoside transporter 1	1	3	0	100

23	P30530	Tyrosine-protein kinase receptor UFO	1	2	0	100
24	P32970	CD70 antigen	1	2	0	100
25	Q8NAV9	cDNA FLJ34690 fis, clone MESAN2000894	1	2	0	100
26	Q8NEY1	Neuron navigator 1	1	2	0	100
27	Q8WWB7	Lysosomal protein NCU-G1	1	2	0	100
28	Q92896	Golgi apparatus protein 1	2	6	1	6
29	P00533	Epidermal growth factor receptor	2	25	6	4.16
30	P48960	CD97 antigen	4	49	12	4.08
31	Q15003	Condensin complex subunit 2	1	12	3	4
32	P24043	Laminin subunit alpha-2	1	22	6	3.66
33	P11117	Lysosomal acid phosphatase	2	10	3	3.33
34	Q30201	Hereditary hemochromatosis protein	1	12	4	3
35	Q9H3G5	Probable serine carboxypeptidase CPVL	1	6	2	3
36	P58743	Prestin	1	3	1	3
37	Q8TCT8	Signal peptide peptidase-like 2A	1	11	4	2.75
38	P07942	Laminin subunit beta-1	3	13	5	2.6
39	Q92945	Far upstream element-binding protein 2	1	13	5	2.6
40	O43852	Calumenin	2	7	3	2.33
41	O60568	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 3	2	8	4	2
42	P04406	Glyceraldehyde-3-phosphate dehydrogenase	1	24	12	2
43	O75976	Carboxypeptidase D	1	6	3	2
44	Q9H6X2	Anthrax toxin receptor 1	1	6	3	2
45	O96005	Cleft lip and palate transmembrane protein 1	1	4	2	2
46	Q08722	Leukocyte surface antigen CD47	1	4	2	2
47	Q13641	Trophoblast glycoprotein	1	4	2	2
48	Q5SZJ2	Heparin sulfate proteoglycan 2	1	2	1	2
49	P14625	Endoplasmic	4	14	29	0.48
50	O00592	Podocalyxin	1	5	11	0.45
51	P07711	Cathepsin L1	2	4	9	0.44
52	Q12797	Aspartyl/asparaginyl beta-hydroxylase	2	3	7	0.42
53	P43308	Translocon-associated protein subunit beta	1	3	7	0.42
54	Q14108	Lysosome membrane protein 2	3	11	26	0.42
55	Q8IWA5	Choline transporter-like protein 2	1	2	5	0.4
56	Q9H6B4	CXADR-like membrane protein	1	2	5	0.4
57	P32004	Neural cell adhesion molecule L1	5	9	23	0.39
58	Q8TCJ2	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3B	2	12	31	0.38
59	P01861	Ig gamma-4 chain C region	2	5	13	0.38
60	P07602	Proactivator polypeptide	7	18	48	0.37
61	P04062	Glucosylceramidase	2	7	19	0.37
62	Q9HDC9	Adipocyte plasma membrane-associated protein	2	4	11	0.36
63	P14314	Glucosidase 2 subunit beta	1	3	9	0.33
64	Q01081	Splicing factor U2AF 35 kDa subunit	1	2	6	0.33
65	Q70UQ0	Inhibitor of nuclear factor kappa-B kinase-interacting protein	1	1	3	0.33
66	P46977	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit STT3A	1	14	44	0.32
67	Q9Y639	Neuroplastin	3	5	17	0.29
68	P11717	Cation-independent mannose-6-phosphate receptor	4	4	14	0.28
69	P15586	N-acetylglucosamine-6-sulfatase	2	4	16	0.25
70	O14773	Tripeptidyl-peptidase 1	1	1	4	0.25
71	O75882	Attractin	1	1	4	0.25
72	O95297	Myelin protein zero-like protein 1	1	1	4	0.25
73	P23229	Integrin alpha-6	1	1	4	0.25
74	Q13421	Mesothelin	3	6	27	0.22
75	Q15165	Serum paraoxonase/arylesterase 2	1	3	14	0.21
76	P26006	Integrin alpha-3	6	4	19	0.21
77	Q5ZPR3	CD276 antigen	3	2	11	0.18
78	P43251	Biotinidase	1	2	11	0.18
79	O00469	Procollagen-lysine,2-oxoglutarate 5-dioxygenase 2	3	2	12	0.17
80	O14672	Disintegrin and metalloproteinase domain-containing protein 10	2	1	10	0.1
81	P05026	Sodium/potassium-transporting ATPase subunit beta-1	1	1	11	0.09

82	Q92542	Nicastrin	5	0	35	0
83	P15328	Folate receptor alpha	4	0	33	0
84	O15031	Plexin-B2	3	0	7	0
85	P10909	Clusterin	2	0	15	0
86	O76082	Solute carrier family 22 member 5	2	0	7	0
87	Q4G148	Glucosidexylosyltransferase 1	2	0	7	0
88	Q99538	Legumain	2	0	7	0
89	P41221	Protein Wnt-5a	2	0	6	0
90	Q99715	Collagen alpha-1(XII) chain	2	0	6	0
91	Q9P273	Teneurin-3	2	0	5	0
92	Q9HAT2	Sialate O-acetyltransferase	2	0	4	0
93	P22309	UDP-glucuronosyltransferase 1-1	1	0	220	0
94	Q96JK2	DDB1- and CUL4-associated factor 5	1	0	35	0
95	P62937	Peptidyl-prolyl cis-trans isomerase A	1	0	32	0
96	Q8NFK8	Torsin-1A-interacting protein 2	1	0	12	0
97	P43007	Neutral amino acid transporter A	1	0	10	0
98	Q5VW38	Protein GPR107	1	0	10	0
99	A8MWK3	Cadherin 2, type 1, N-cadherin (neuronal)	1	0	8	0
100	P05362	Intercellular adhesion molecule 1	1	0	8	0
101	Q6P4E1	Protein CASC4	1	0	8	0
102	Q8N766	Uncharacterized protein KIAA0090	1	0	8	0
103	Q8TEM1	Nuclear pore membrane glycoprotein 210	1	0	8	0
104	Q9H0X4	Protein ITFG3	1	0	8	0
105	P08842	Steryl-sulfatase	1	0	7	0
106	P07954	Fumarate hydratase, mitochondrial	1	0	6	0
107	P08648	Integrin alpha-5	1	0	6	0
108	Q32P28	Prolyl 3-hydroxylase 1	1	0	6	0
109	Q68CQ7	Glycosyltransferase 8 domain-containing protein 1	1	0	6	0
110	Q9H330	Transmembrane protein C9orf5	1	0	6	0
111	Q8WXI7	Mucin-16	1	0	5	0
112	O00468	Agrin	1	0	4	0
113	O95857	Tetraspanin-13	1	0	4	0
114	P16278	Beta-galactosidase	1	0	4	0
115	Q02388	Collagen alpha-1(VII) chain	1	0	4	0
116	Q92626	Peroxidase homolog	1	0	4	0
117	Q9BY67	Cell adhesion molecule 1	1	0	4	0
118	B7Z553	cDNA FLJ51266, highly similar to Vitronectin	1	0	3	0
119	E9PQH3	Zinc finger protein 585B	1	0	3	0
120	F5GYS1	Protein phosphatase 2, regulatory subunit B', delta	1	0	3	0
121	P08473	Nephrilysin	1	0	3	0
122	P11166	Solute carrier family 2, facilitated glucose transporter member 1	1	0	3	0
123	P12259	Coagulation factor V	1	0	3	0
124	P13674	Prolyl 4-hydroxylase subunit alpha-1	1	0	3	0
125	P19075	Tetraspanin-8	1	0	3	0
126	P21589	5'-nucleotidase	1	0	3	0
127	P35052	Glypican-1	1	0	3	0
128	P48307	Tissue factor pathway inhibitor 2	1	0	3	0
129	P55058	Phospholipid transfer protein	1	0	3	0
130	Q9BVX2	Transmembrane protein 106C	1	0	3	0
131	Q9NXH8	Torsin-4A	1	0	3	0
132	D6RA51	Toll-like receptor 3	1	0	2	0
133	P54709	Sodium/potassium-transporting ATPase subunit beta-3	1	0	2	0
134	P55268	Laminin subunit beta-2	1	0	2	0
135	Q13332	Receptor-type tyrosine-protein phosphatase S	1	0	2	0
136	Q13443	Disintegrin and metalloproteinase domain-containing protein 9	1	0	2	0
137	Q13586	Stromal interaction molecule 1	1	0	2	0
138	Q5JRA6	Melanoma inhibitory activity protein 3	1	0	2	0
139	Q86SE7	Tissue factor	1	0	2	0
140	Q8N697	Solute carrier family 15 member 4	1	0	2	0

141	Q8NA58	Poly(A)-specific ribonuclease PARN-like domain-containing protein 1	1	0	2	0
142	Q96K49	Transmembrane protein 87B	1	0	2	0
143	Q9P2B2	Prostaglandin F2 receptor negative regulator	1	0	2	0
144	Q9Y2C2	Uronyl 2-sulfotransferase	1	0	2	0

Proteins with spectral count ratio above 2 and below 0.2 were shown here. For protein uniquely identified from one sample, at least two matched spectra were required to report here. Please see Table S2 for the complete list.

Column A shows the index of identified glycopeptides

Column B shows the IPI accession of identified glycoproteins.

Column C shows the Uniprot Accession of identified glycoproteins.

Column D shows the names of identified glycoproteins.

Column E shows the number of glycopeptide identified from each glycoprotein.

Column F and G shows the combined spectra counts for each protein in HepG2 sample and L02 sample, respectively.

Column H shows spectra ratio of each glycoprotein between HepG2 and L02 cells. For the proteins only identified from HepG2 cells, an arbitrarily number 100 is assigned.

Table 2: Label-free quantification of identified glycoproteins (partial).

Epidermal growth factor receptor (EGFR) is up-regulated in HepG2 cells than in L02 cells with a spectral count ratio above 4. This observation is in agree with Hoffmann et al. showing EGF activation enhances the proliferation of resistant cancer cell [51]. Beyond the several examples shown above, Golgi apparatus protein 1 (GLG1) is another interesting target which is up-regulated in ovarian cancer [52].

Acknowledgment

This work is supported by NFSC (Grant No. 81372365) and Foundation of Shaanxi Educational Committee (Grant No. 12JK0830).

References

- Raney B (1997) Hepatoblastoma in Children: A Review. *J Pediatr Hematol Oncol* 19: 418-422.
- Ishak KG, Glunz PR (1967) Hepatoblastoma and hepatocarcinoma in infancy and childhood. Report of 47 cases. *Cancer* 20: 396-422.
- Choufani S, Shuman C, Weksberg R (2010) Beckwith–Wiedemann syndrome. *Am J Med Genet C Semin Med Genet* 154: 343-354.
- Galiatsatos P, Foulkes WD (2006) Familial Adenomatous Polyposis. *Am J Gastroenterol* 101: 385-398.
- Luo J-H, Ren B, Keryanov S, Tseng GC, Rao UNM, et al. (2006) Transcriptomic and genomic analysis of human hepatocellular carcinomas and hepatoblastomas. *Hepatology* 44: 1012-1024.
- Blazquez M, Carretero A, Ellis JK, Athersuch TJ, Cavill R, et al. (2013) A Combination of Transcriptomics and Metabolomics Uncovers Enhanced Bile Acid Biosynthesis in HepG2 Cells Expressing CCAAT; Enhancer-Binding Protein β (C/EBP β), Hepatocyte Nuclear Factor 4 α (HNF4 α), and Constitutive Androstane Receptor (CAR). *J Proteome Res* 12: 2732-2741.
- Pattanakisakul S-n, Rungrojcharoenkit K, Kanlaya R, Sinchaikul S, Noisakran S, et al. (2007) Proteomic Analysis of Host Responses in HepG2 Cells during Dengue Virus Infection. *J Proteome Res* 6: 4592-4600.
- Tong A, Gou L, Lau QC, Chen B, Zhao X, et al. (2008) Proteomic Profiling Identifies Aberrant Epigenetic Modifications Induced by Hepatitis B Virus X Protein. *J Proteome Res* 8: 1037-1046.
- Hagglund P, Bunkenborg J, Elortza F, Jensen ON, Roepstorff P (2004) A New Strategy for Identification of N-Glycosylated Proteins and Unambiguous Assignment of Their Glycosylation Sites Using HILIC Enrichment and Partial Deglycosylation. *J Proteome Res* 3: 556-566.
- Helenius A, Aebi M (2001) Intracellular Functions of N-Linked Glycans. *Science* 291: 2364-2369.
- Rudd PM, Elliott T, Cresswell P, Wilson IA, Dwek RA (2001) Glycosylation and the Immune System. *Science* 291: 2370-2376.
- Ohtsubo K, Marth JD (2006) Glycosylation in Cellular Mechanisms of Health and Disease. *Cell* 126: 855-867.
- Freeze HH, Aebi M (2005) Altered glycan structures: the molecular basis of congenital disorders of glycosylation. *Curr Opin Struct Biol* 15: 490-498.
- Leroy JG (2006) Congenital Disorders of N-Glycosylation Including Diseases Associated With O- as Well as N-Glycosylation Defects. *Pediatr Res* 60: 643-656.
- Durand G, Seta N (2000) Protein Glycosylation and Diseases: Blood and Urinary Oligosaccharides as Markers for Diagnosis and Therapeutic Monitoring. *Clin Chem* 46: 795-805.
- Miyamoto S (2006) Clinical applications of glycomic approaches for the detection of cancer and other diseases. *Curr Opin Mol Ther* 8: 507-513.
- Roth J (2002) Protein N-glycosylation along the secretory pathway: relationship to organelle topography and function, protein quality control, and cell interactions. *Chem Rev* 102: 285-303.
- Kaji H, Saito H, Yamauchi Y, Shinkawa T, Taoka M, et al. (2003) Lectin affinity capture, isotope-coded tagging and mass spectrometry to identify N-linked glycoproteins. *Nat Biotechnol* 21: 667-672.
- Kaji H, Yamauchi Y, Takahashi N, Isobe T (2007) Mass spectrometric identification of N-linked glycopeptides using lectin-mediated affinity capture and glycosylation site-specific stable isotope tagging. *Nat Protoc* 1: 3019-3027.
- Sturiale L, Barone R, Palmigiano A, Ndosimao CN, Briones P, et al. (2008) Multiplexed glycoproteomic analysis of glycosylation disorders by sequential yolk immunoglobulin immunoseparation and MALDI-TOF MS. *Proteomics* 8: 3822-3832.
- Cho W, Jung K, Regnier FE (2008) Use of Glycan Targeting Antibodies to Identify Cancer-Associated Glycoproteins in Plasma of Breast Cancer Patients. *Anal Chem* 80: 5286-5292.
- Sparbier K, Asperger A, Resemann A, Kessler I, Koch S, et al. (2007) Analysis of Glycoproteins in Human Serum by Means of Glycospecific Magnetic Bead Separation and LC-MALDI-TOF; TOF Analysis with Automated Glycopeptide Detection. *J Biomol Tech* 18: 252-258.
- Xu Y, Wu Z, Zhang L, Lu H, Yang P, et al. (2008) Highly Specific Enrichment of Glycopeptides Using Boronic Acid-Functionalized Mesoporous Silica. *Anal Chem* 81: 503-508.
- Alvarez-Manilla G, Atwood, Guo Y, Warren NL, Orlando R, et al. (2006) Tools for Glycoproteomic Analysis: Size Exclusion Chromatography Facilitates Identification of Tryptic Glycopeptides with N-linked Glycosylation Sites. *J Proteome Res* 5: 701-708.
- Jia W, Lu Z, Fu Y, Wang H-P, Wang L-H, et al. (2009) A Strategy for Precise and Large Scale Identification of Core Fucosylated Glycoproteins. *Mol Cell Proteomics* 8: 913-923.
- Lewandrowski U, Lohrig K, Zahedi R, Wolters D, Sickmann A (2008) Glycosylation Site Analysis of Human Platelets by Electrostatic Repulsion Hydrophilic Interaction Chromatography. *Clin Proteomics* 4: 25-36.
- Zhang H, Guo T, Li X, Datta A, Park JE, et al. (2010) Simultaneous Characterization of Glyco- and Phosphoproteomes of Mouse Brain Membrane Proteome with Electrostatic Repulsion Hydrophilic Interaction Chromatography. *Mol Cell Proteomics* 9: 635-647.
- Tajiri M, Yoshida S, Wada Y (2005) Differential analysis of site-specific glycan on plasma and cellular fibronectins: application of a hydrophilic affinity method for glycopeptide enrichment. *Glycobiology* 15: 1332-1340.

29. Zhang H, Li XJ, Martin DB, Aebersold R (2003) Identification and quantification of N-linked glycoproteins using hydrazide chemistry, stable isotope labeling and mass spectrometry. *Nat Biotechnol* 21: 660-666.
30. Tian YA, Zhou Y, Elliott S, Aebersold R, Zhang H (2007) Solid-phase extraction of N-linked glycopeptides. *Nat Protoc* 2: 334-339.
31. Mysling S, Palmisano G, Hojrup P, Thaysen-Andersen M (2010) Utilizing Ion-Pairing Hydrophilic Interaction Chromatography Solid Phase Extraction for Efficient Glycopeptide Enrichment in Glycoproteomics. *Anal Chem* 82: 5598-5609.
32. Kuroguchi M, Matsushita T, Amano M, Furukawa J-i, Shinohara Y, et al. (2010) Sialic Acid-focused Quantitative Mouse Serum Glycoproteomics by Multiple Reaction Monitoring Assay. *Mol Cell Proteomics* 9: 2354-2368.
33. Cao J, Shen C, Wang H, Shen H, Chen Y, et al. (2009) Identification of N-Glycosylation Sites on Secreted Proteins of Human Hepatocellular Carcinoma Cells with a Complementary Proteomics Approach. *J Proteome Res* 8: 662-672.
34. Crooks GE, Hon G, Chandonia JM, Brenner SE (2004) Web Logo: A Sequence Logo Generator. *Genome Res* 14: 1188-1190.
35. Tian Y, Esteva FJ, Song J, Zhang H (2012) Altered Expression of Sialylated Glycoproteins in Breast Cancer Using Hydrazide Chemistry and Mass Spectrometry. *Mol Cell Proteomics* 11: M111.011403.
36. Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, et al. (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21: 3674-3676.
37. Consortium TU (2008) The Universal Protein Resource (UniProt). *Nucleic Acids Res* 36: 190-195.
38. Sethi MK, Buettner FFR, Krylov VB, Takeuchi H, Nifantiev NE, et al. (2010) Identification of Glycosyltransferase 8 Family Members as Xylosyl transferases Acting on O-Glycosylated Notch Epidermal Growth Factor Repeats. *J Biol Chem* 285: 1582-1586.
39. Chen R, Jiang X, Sun D, Han G, Wang F, et al. (2009) Glycoproteomics Analysis of Human Liver Tissue by Combination of Multiple Enzyme Digestion and Hydrazide Chemistry. *J Proteome Res* 8: 651-661.
40. Udomuksorn W, Elliot DJ, Lewis BC, Mackenzie PI, Yoovathaworn K, et al. (2007) Influence of mutations associated with Gilbert and Crigler-Najjar type II syndromes on the glucuronidation kinetics of bilirubin and other UDP-glucuronosyltransferase 1A substrates. *Pharmacogenet Genomics* 17: 1017-1029.
41. Hanover JA (2001) Glycan-dependent signaling: O-linked N-acetyl glucosamine. *The FASEB Journal* 15: 1865-1876.
42. Turner JR, Tartakoff AM, Greenspan NS (1990) Cytologic assessment of nuclear and cytoplasmic O-linked N-acetyl glucosamine distribution by using anti-streptococcal monoclonal antibodies. *Proc Natl Acad Sci U S A* 87: 5608-5612.
43. Shakin-Eshleman SH, Spitalnik SL, Kasturi L (1996) The Amino Acid at the X Position of an Asn-X-Ser Sequon Is an Important Determinant of N-Linked Core-glycosylation Efficiency. *J Biol Chem* 271: 6363-6366.
44. Mellquist JL, Kasturi L, Spitalnik SL, Shakin-Eshleman SH (1998) The Amino Acid Following an Asn-X-Ser,Thr Sequon Is an Important Determinant of N-Linked Core Glycosylation Efficiency. *Biochemistry* 37: 6833-6837.
45. Bause E, Legler G (1981) The role of the hydroxy amino acid in the triplet sequence Asn-Xaa-Thr (Ser) for the N-glycosylation step during glycoprotein biosynthesis. *Biochem J* 195: 639-644.
46. Chen R, Tan Y, Wang M, Wang F, Yao Z, et al. (2011) Development of Glycoprotein Capture-Based Label-Free Method for the High-throughput Screening of Differential Glycoproteins in Hepatocellular Carcinoma. *Mol Cell Proteomics* 10: M110.006445.
47. Neilson KA, Ali NA, Muralidharan S, Mirzaei M, Mariani M, et al. (2011) Less label, more free: Approaches in label-free quantitative mass spectrometry. *Proteomics* 11: 535-553.
48. Liu H, Sadygov RG, Yates JR (2004) A Model for Random Sampling and Estimation of Relative Protein Abundance in Shotgun Proteomics. *Anal Chem* 76: 4193-4201.
49. Golubkov VS, Chekanov AV, Cieplak P, Aleshin AE, Chernov AV, et al. (2010) The Wnt/Planar Cell Polarity Protein-tyrosine Kinase-7 (PTK7) Is a Highly Efficient Proteolytic Target of Membrane Type-1 Matrix Metalloproteinase: IMPLICATIONS IN CANCER AND EMBRYOGENESIS. *J Biol Chem* 285: 35740-35749.
50. Lombardo Y, Filipović A, Molyneux G, Periyasamy M, Giamas G, et al. (2012) Nicastrin regulates breast cancer stem cell properties and tumor growth in vitro and in vivo. *Proceedings of the National Academy of Sciences* 109: 16558-16563.
51. Hoffmann K, Xiao Z, Franz C, Mohr E, Serba S, et al. (2011) Involvement of the epidermal growth factor receptor in the modulation of multidrug resistance in human hepatocellular carcinoma cells in vitro. *Cancer Cell International* 11: 40.
52. Pan S, Cheng L, White JT, Lu W, Utleg AG, et al. (2009) Quantitative Proteomics Analysis Integrated with Microarray Data Reveals That Extracellular Matrix Protein, Catenins, and P53 Binding Protein 1 Are Important for Chemotherapy Response in Ovarian Cancers. *OMICS: A Journal of Integrative Biology* 13: 345-354.