Published in final edited form as:

J Proteome Res. 2015 November 6; 14(11): 4876–4884. doi:10.1021/acs.jproteome.5b00718.

Mass-Selected Site-Specific Core-Fucosylation of Serum Proteins in Hepatocellular Carcinoma

Haidi Yin[†], Zhijing Tan[†], Jing Wu[†], Jianhui Zhu[†], Kerby A. Shedden[‡], Jorge Marrero[§], and David M. Lubman^{†,*}

[†]Department of Surgery, University of Michigan Medical Center, Ann Arbor, Michigan 48109, United States

[‡]Department of Statistics, University of Michigan, Ann Arbor, Michigan 48109, United States

[§]Liver Transplantation Program, University of Texas Southwestern Medical Center, Dallas, Texas 75390, United States

Abstract

A mass spectrometry-based methodology has been developed to screen for changes in site-specific core-fucosylation (CF) of serum proteins in early stage HCC with different etiologies. The methods involve depletion of high-abundance proteins, trypsin digestion of medium-to-low-abundance proteins into peptides, iTRAQ labeling, and *Lens culinaris* Agglutinin (LCA) enrichment of CF peptides, followed by endoglycosidase F3 digestion before mass spectrometry analysis. 1300 CF peptides from 613 CF proteins were identified from patients sera, where 20 CF peptides were differentially expressed in alcohol (ALC)-related HCC samples compared with ALC-related cirrhosis samples and 26 CF peptides changed in hepatitis C virus (HCV)-related HCC samples compared with HCV-related cirrhosis samples. Among these, we found three CF peptides from fibronectin upregulated in ALC-related HCC samples compared with ALC-related cirrhosis samples with an AUC (area under the curve) value of 0.89 at site 1007 with a specificity of 85.7% at a sensitivity of 92.9% (generated with 10-fold cross-validation). When combined with the AFP index, the AUC value reached to 0.92 with a specificity of 92.9% at a sensitivity of 100%, significantly improved compared to that with AFP alone (LR test $p < 0.001$). In HCV-related samples, the CF level of cadherin-5 at site 61 showed the best AUC value of 0.75 but was not as promising as that of fibronectin site 1007 for ALC-related samples. The CF peptides of fibronectin may serve as potential biomarkers for early stage HCC screening in ALC-related cirrhosis patients.

Graphical abstract

*Corresponding Author: dmlubman@umich.edu; Phone: 734-647-8834; Fax: 734-615-2088.

The authors declare no competing financial interest.

Keywords

Core-fucosylation; hepatocellular carcinoma; site-specific; alcohol-related HCC; hepatitis C virus-related HCC

INTRODUCTION

Hepatocellular carcinoma (HCC) is the third most common cause of cancer-related death worldwide.¹ 80 to 90% of HCC develops in patients that already have liver cirrhosis² and have been recommended for surveillance for an early onset of HCC. In the USA, Europe, and Japan, treatments can be applied only to 30% of the patients who are diagnosed at early stage, resulting in a 5 year survival rate higher than 50%, whereas the patients diagnosed at late stage have a 2 year survival rate lower than 16%.³ Therefore, diagnosis of HCC at early stage is much needed and critical for survival.

The most common noninvasive detection of HCC involves imaging (ultrasound, computer tomography, and magnetic resonance imaging). Serum markers are used as a complementary method in patients with cirrhosis or small tumors.² Serum alpha-fetoprotein (AFP) is most widely used as a clinical HCC diagnostic marker; however, AFP is not very sensitive. A study has shown that AFP has a specificity of 90.6% and a sensitivity of 60.0% at the cutoff value of 20 ng/mL.⁴ Desgamma carboxy prothrombin (DCP) is widely used as an alternative marker for AFP in Japan. However, the diagnosis value of DCP varies among different groups of patients. In patients with viral etiology, DCP has an AUC value of 0.76, whereas in patients with nonviral etiology, it has an AUC value of 0.65.⁵

Along with seeking markers that change at the protein level, new strategies are needed to improve HCC early detection. Protein post-transcriptional modification, especially glycosylation,^{6–8} has also been shown to change in different disease states. In particular, core-fucosylation (CF) recently has been used as a potential marker for various cancers, including pancreatic cancer,^{9–11} lung cancer,¹² and liver cancer.^{13–16} AFP-L3, where *Lens culinaris* Agglutinin (LCA) binds to the core-fucosylated (CF) glycopeptide of AFP, has been used as an alternative marker for HCC.¹⁵ In Japan AFP-L3 has been developed as a diagnostic kit for patients with AFP < 20 ng/ mL, where the AUC value between HCC and chronic liver disease was 0.707, with a specificity of 85.1% at a sensitivity of 41.5%.¹⁶

The three most common etiologies of the underlying liver disease in HCC are ALC (alcohol)-related, HCV (hepatitis C virus)-related, and HBV (hepatitis B virus)-related. Another rising etiology is nonalcoholic fatty liver disease. Currently used biomarkers for HCC, including AFP, DCP, and AFP-L3, combine all cirrhosis or HCC samples of various etiologies for analysis. Biomarker screening strategies tailored for each etiology may prove to be better at detection of early stage HCC. In previous work, ceruloplasmin was found to be upregulated in LCA-enriched HCC serum samples, indicating the upregulation of CF ceruloplasmin.¹⁷ A more detailed work has shown that the CF ratio of ceruloplasmin at site 138 provides an AUC value of 0.922 between normal and ALC-related cirrhosis and an AUC value of 0.838 between ALC-related cirrhosis and ALC-related HCC. However, this change was not found in patients with HBV or HCV etiology related patients.¹⁸

In this work, we investigated the total serum screening of all of the CF peptides in serum from stage A early stage HCC and cirrhosis samples of ALC and HCV etiologies. The CF levels were compared between disease states of each etiology. ALC and HCV are the two major causes of HCC in the USA, whereas HBV is the major cause in Asia. Clinical samples used in this study were collected in the USA so that only ALC- and HCV-related samples were examined.

MATERIALS AND METHODS

Serum Samples

All of the serum samples were collected at the University of Michigan Hospital. HCC samples were stage A early HCC according to the Barcelona Clinical Liver Cancer (BCLC) staging system. 52 serum samples were used in this study, composed of 14 ALC-related cirrhosis samples, 14 ALC-related early stage HCC samples, 12 HCV-related cirrhosis samples, and 12 HCV-related early stage HCC samples. This study was approved by the Institutional Review Board of the University of Michigan Medical School. All serum samples were stored at -80°C before use. Basic clinical information is given in Table 1. All clinical investigations were conducted according to the principles expressed in the Declaration of Helsinki.

Depletion of High-Abundance Proteins

Serum top 10 high-abundance proteins were depleted using an IgY14 LC5 column (Sigma, St. Louis). The buffer of the eluted sample was changed to 50 mM TEAB with a 3k Ultra centrifugal filter-15 (Millipore Amicon) at 5000g for 1 h, which was repeated three times. The processed sample was then stored at -20°C .

Trypsin Digestion and iTRAQ Labeling

100 μg of depleted protein was reduced by 25 mM TCEP in 4 M urea at 37°C for 1 h followed by alkylation with 20 mM IAA in the dark for 30 min. The buffer of the mixture was changed to 50 mM TEAB with a 3k Ultra centrifugal filter-15 at 5000g for 1 h, and the process was repeated twice. Trypsin (1:30) was added to the digestion system at 37°C overnight. The digested sample was dried by SpeedVac (Labconco, Kansas City, MO). An aliquot of each sample was pooled as an internal standard. iTRAQ reagent was used for

quantitative analysis, where one set includes one cirrhosis sample, one HCC sample, and one internal standard. The 4 iTRAQ tags were randomly assigned to the three samples in a set and kept at room temperature for 1 h followed by quenching with 30 μ L of 5% $\text{NH}_3\cdot\text{H}_2\text{O}$ for 15 min. The three samples were afterward pooled and dried using a SpeedVac. The labeled samples were then buffer-exchanged to LCA binding buffer (20 mM Tris, 0.15 M NaCl, pH7.4, 1 mM CaCl_2 , 1 mM MnCl_2) with a 3k Ultra centrifugal filter-4 at 7500g for 1 h, and the process was repeated twice.

LCA Enrichment

Nonglycosylated peptides severely affect the ionization efficiency of glycosylated peptides in the electrospray process. An LCA column¹⁹ was used, therefore, to enrich the CF peptides. A pooled iTRAQ-labeled sample (100 μ g of HCC, 100 μ g of cirrhosis, and 100 μ g of internal standard) was loaded on 1 mL of LCA resin in a 2 mL column (Fisher Scientific, Pittsburgh) after washing the resin with 2 mL of LCA binding buffer three times. Incubation was performed at room temperature for 15 min with gentle rotation, and the solution was eluted and reloaded onto the resin again. After another 15 min of incubation, the resin was washed by 4 mL of binding buffer four times. The bound CF peptides were eluted by 800 μ L of elution buffer A (binding buffer with 200 mM mannoside and 200 mM glucoside) twice and by 800 μ L of elution buffer B (20 mM Tris with 200 mM mannoside and 200 mM glucoside).¹⁰ All eluent was combined for buffer exchange to 50 mM sodium acetate with a 3k Ultra centrifugal filter-4 (Millipore Amicon) at 7500g for 1 h, and the process was repeated twice.

Endo F3 Digestion

Endo F3 2 μ L (10 mU) (QA-Bio, San Mateo, CA) was added to the buffer-exchanged sample for overnight digestion at 37 °C. Endo F3 digested samples were deactivated at 95 °C for 5 min, and TFA was added to make a final concentration at 0.1%. The samples were desalted with a C18 column with modification. A C18 column was activated with 200 μ L of activation solution (50% ACN with 0.1% TFA) five times and washed with 200 μ L of equilibration solution (0.1% TFA) three times by centrifugation at 1500g for 1 min. The sample was loaded on top of the resin, and the eluent was collected after centrifugation at 1500g for 1 min. The eluent was loaded on top of the resin again, and the above step was repeated four times. The resin was then washed by 200 μ L of equilibration solution five times, followed by elution with 20 μ L of elution buffer A (50% ACN, 0.1% TFA) and 20 μ L of elution buffer B (75% ACN, 0.1% TFA), which was performed twice with each elution buffer. All eluents were mixed and aliquoted into four tubes. The sample was dried in a SpeedVac and stored at -20 °C before analysis by LC-MS.

Nano LC-Q Exactive MS

An Acclaim PepMap RSLC C18 column (75 μ m i.d. $\times$ 25 cm; 2 μ m particles, 100 Å, nanoViper) was used for LC separation, and gradient elution was performed using an EASY-nLC 1000 liquid chromatograph system (Thermo Fisher Scientific, San Jose, CA) with a flow rate at 300 nL/min. Mobile phase A was 2% acetonitrile with 0.1% formic acid in water, and mobile phase B was 2% water with 0.1% formic acid in acetonitrile. The

analytical gradient lasted for 80 min, where the composition of solvent B changed from 5 to 35%, followed by a washing and equilibration step where solvent B increased to 95% in 5 min, was held for 2 min, and then returned to 2% B in 2 min and was held for 8 min.

An ESI-Q Exactive Hybrid Quadrupole-Orbitrap mass spectrometer (Thermo Fisher Scientific, San Jose, CA) operated in positive ion mode was used for analysis. The ESI spray voltage was set at 2.5 kV. High-energy collision dissociation (HCD) fragmentation was performed at 32% of the normalized collision energy. The mass spectra were acquired in a data-dependent manner. Following a full scan in the mass range of m/z 350 to 1600, HCD MS/MS was performed on the first to the fourth most intense ions from the survey MS full scan.

Database Search

The raw data was searched with Proteome Discoverer1.4 (Thermo Fisher Scientific, San Jose, CA) software with SEQUEST using the following settings: (1) one missed cleavage was allowed; (2) cysteine carbamidomethylation (+57.021 Da) and lysine iTRAQ labeling (+144.102) were set as fixed modifications; (3) methionine oxidation (+15.995 Da) and addition of GlcNAc + Fucose (+349.137 Da) to an asparagine residue were set as dynamic modifications; (4) peptide ion tolerance of 10 ppm and fragment ion tolerance of 0.05 Da; and (6) SWISS-PROT *Homo sapiens* database (reviewed, downloaded in April, 2014) was used. Peptide matches that pass the filter associated with FDR < 0.01 were used for analysis.

Statistical Analysis

The CF level of each site of serum proteins in different liver diseases was compared using GraphPad Prism 5 (GraphPad, La Jolla, CA). D'Agostino–Pearson omnibus normality test was used to check the distribution of CF levels at each site before the following Student's *t*-test. For normally distributed groups, the difference between means was analyzed using the *t*-test with Welch's correction at confidence intervals of 95% and two-tailed *p* values; for non-normal distributed groups, the difference between means was analyzed using the Mann–Whitney test at confidence intervals of 95% and two-tailed *p* values. The threshold $p < 0.05$ was used to determine the differentially expressed CF peptides. The receiver operator characteristic (ROC) analysis was performed by 10-fold cross-validation using Python considering the relatively small sample size. The AFP values were log-transformed before analysis.

RESULTS AND DISCUSSION

Most developed methods for CF protein study use lectin enrichment at the protein level and/or cut off the entire glycan structure with PNGase F prior to CF proteins identification with mass spectrometry.^{11,12,24} These studies provide neither information about the CF site nor the CF level of each specific CF site. In the procedure used in this study, lectin enrichment was performed at the peptide level and the glycan was truncated partially for CF identification so that not only the CF site but also the CF level of each specific CF site was analyzed. For CF level quantification, there are two methods: one is label-free relative

quantification^{11,18} and the other one is absolute quantification with iTRAQ labeling.¹⁰ iTRAQ labeling was used to quantify the CF peptides in this work.

Previous work has shown that the CF ratio of ceruloplasmin at site 138 is a possible marker to distinguish ALC-related HCC from ALC-related cirrhosis, whereas in patients with HBV or HCV etiology, no significant change was observed in cirrhosis samples versus HCC samples.¹⁸ In this study, we screened site-specific CF changes of serum glycoproteins in liver cirrhosis and HCC patients with two etiologies, including ALC and HCV, in order to search for additional specific changes that can be used as markers of HCC. The workflow is shown in Figure S1.

CF Site Identification

Each sample was run on the Q Exactive mass spectrometer three times, based on the fact that when one sample was run three times 91% of the peptides was identified in two runs (Figure 1A); when one sample was run four times, 94% was identified in three runs, whereas 86% was identified in two runs, as shown in Figure 1B. Around 400–500 CF peptides corresponding to 150–300 CF proteins were identified in each sample with a combination of three runs. The list of CF peptides identified is shown in Table S1. 1300 CF peptides from 613 proteins were identified in this study. The percentage of CF peptides identified in the number of sets is shown in Figure S2. 833 (64.1%) identified CF peptides were observed in 4 sets out of the 26 sets, where 556 (42.8%) identified CF peptides were observed in 10 sets. With the procedure used in this study, nine sets of samples were found to be sufficient to identify more than 95% of the total identified CF peptides in either ALC-related or HCV-related samples (Figure 2). Therefore, 14 sets of ALC-related samples and 12 sets of HCV-related samples were sufficient to identify most of the CF serum proteins. In this study, a Q Exactive MS was applied. The spectra of CF peptides identified by Q Exactive MS are similar to that identified by Orbitrap Elite MS but with improved signal intensity (Figure S3). The spectra of peptides with both single CF and double CF have good fragmentation using Q Exactive MS.

Around 40% of the identified CF proteins localize at the extracellular space, another 40% at the plasma membrane, 10% at the cytoplasm, 2% at the nucleus, and 2% at other localizations in both ALC-related and HCV-related samples (Figure S4). Slight differences were found between ALC-related and HCV-related samples: less CF proteins located at plasma membrane (38.42%) were found in ALC-related samples than in HCV-related samples (44.83%), whereas more CF proteins located at extracellular space (46.45%) were found in ALC-related samples than in HCV-related samples (38.66%), indicating that alcohol intake may promote the core-fucosylation of proteins in the extracellular space,

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jproteome.5b00718](https://doi.org/10.1021/acs.jproteome.5b00718).

List of all CF peptides identified in this study (Table S1) (XLSX).

Workflow of this study (Figure S1); percentage of CF peptides identified in the number of sets (Figure S2); spectra of identified glycopeptides with 1 and 2 CF sites in Q exactive compared with Orbitrap Elite (Figure S3); localization of identified CF proteins (Figure S4); analysis of the CF level of GP73 at site 109 (Figure S5); analysis of the CF level of hemopexin at sites 187 and 453 (Figure S6); analysis of the CF level of cadherin-5 at site 61 (Figure S7); CF level of some other differentially expressed CF peptides in HCC samples compared with cirrhosis samples ALC-related samples HCV-related samples (Figure S8) (PDF).

whereas HCV infection may promote the core-fucosylation of membrane proteins or the shedding of membrane CF proteins.

Differentially Expressed CF Peptides in ALC-Related HCC Compared to ALC-Related Cirrhosis

ALC-related cirrhosis accounts for one-third to one-half of the cases of HCC in nonvirus endemic areas.²⁰ In the ALC-related samples, 15 CF peptides from 10 CF proteins were found to be upregulated in HCC compared to cirrhosis, whereas five CF peptides from five CF proteins were downregulated (Table 2). Differentially expressed CF peptides identified in less than six sets were excluded. These 15 CF proteins were analyzed with IPA (Ingenuity Pathway Analysis, Qiagen). The top network involves nine CF proteins: clusterin (CLU), dopamine beta-hydroxylase (DBH), coagulation factor VIII (F8), fibronectin 1 (FN1), intercellular adhesion molecule 3 (ICAM3), insulin-like growth factor binding protein 3 (IGFBP3), thrombospondin 1 (THBS1), von Willebrand factor (VWF), and prothrombin (F2). These proteins are involved interactively in cell-to-cell signaling and interaction, hematological system development and function, and immune cell trafficking (Figure 3A).

Among these changed CF peptides, a CF peptide (aVLVNnITTGER) of GP 73 with site 109 was found to be upregulated in HCC samples compared to cirrhosis samples ($p = 0.029$) (the small letters refer to the residues being modified: N-terminal is iTRAQ labeled; n indicates core-fucosylated asparagine). GP73 (Golgi membrane protein 73) is a Golgi resident glycoprotein that normally is present at low abundance (nanogram levels) in the circulation of healthy subjects and has been found to be correlated with the presence of HCC in many studies.^{21–23} CF GP73 has also been proposed to be a marker in a study with small sample size, but no study on detailed site-specific quantification of CF GP73 has been undertaken so far.²⁴ The AUC value for CF GP73 was 0.73 with a specificity of 72.7% at sensitivity of 81.8% vs 0.68 for AFP in our study with a combined ROC analysis of 0.81 (Figure S5). From the scatter plot, it is apparent that some cirrhosis samples also had high CF GP73 values, similar to GP73. The high expression level of GP73 in hepatitis and cirrhosis has been discussed by Iftikhar and co-workers.²⁵

Fibronectin was one of the most significantly changed CF proteins in this study. Fibronectin has theoretically seven N-glycosylation sites: sites 430, 528, 542, 877, 1007, 1244, and 2108. This study detected three CF sites, including sites 528 (CF peptide: dQcIVDDITYNVnDTFhk), 542 (CF peptide: hEEGHmLncTcFGQGR), and 1007 (CF peptide: IDAPTNLQFVnETDSTVLVR) (the small letters refer to the residues being modified: N-terminal is iTRAQ labeled; m indicates methionine oxidation; n indicates core-fucosylated asparagine; c indicates carbamidomethylation of cysteine.). The expression level of the CF peptides with these three sites increased in ALC-related HCC samples compared to cirrhosis samples, and this increase was found only in ALC-related samples, not in HCV-related samples. The quantitative comparison and the ROC analysis are shown in Figure 4. All three sites showed very high AUC values, with site 528 at 0.78 (with a specificity of 78.6% at a sensitivity of 85.7%), site 542 at 0.87 (with a specificity of 85.7% at a sensitivity of 92.9%), and site 1007 at 0.89 (with a specificity of 85.7% at a sensitivity of 92.9%). A combined analysis scatter plot with AFP is shown in Figure 5. When combined with the

clinical AFP value, the AUC value for site 1007 between ALC-related cirrhosis and HCC increased to 0.92 with a specificity of 92.9% at a sensitivity of 100% (Figure 5C), whereas for sites 528 and 542, the combined AUC values were 0.81 and 0.87, respectively (Figure 5A,B) (generated by 10-fold cross-validation). The model fit of site 1007 was significantly improved over the one with AFP alone (AUC value = 0.67) (LR test $p < 0.001$). Both sites 542 and 1007 can well-distinguish ALC-related HCC from cirrhosis samples, whereas site 1007 had significantly improved specificity and sensitivity. The CF level of site 1007 of fibronectin was elevated in the serum of 70% (7/10) of the patients with HCC who had serum AFP levels less than 10 ng/mL (Figure 5C). DCP is used as a marker in Japan to complement AFP and also has shown good results in American patients.²⁶ However, DCP does not have good performance in patients with nonviral etiology.⁵ Our previous work showed that the relative CF ratio of ceruloplasmin at site 138 has an AUC level of 0.836 with a specificity of 77.8% and a sensitivity of 79.2% in ALC-related samples.¹⁸ Here, the CF peptide level of fibronectin at site 1007 provides even better diagnostic value.

Plasma fibronectin (soluble dimeric form) is secreted by hepatocytes.²⁷ Core fucose is usually added to the GlcNAc in medial Golgi from GDP-fucose by fucose transferase VIII (FucT-VIII).²⁸ The upregulation of the CF level of fibronectin is, therefore, dependent on the level of FucT-VIII in hepatocytes. FucT-VIII of hepatocytes has been investigated, and the results showed the upregulation of FucT-VIII in HCC patients.²⁹ The level of fibronectin in serum, therefore, may indicate the level of FucT-VIII in hepatocytes. The value of the CF level of fibronectin at site 1007 may complement DCP in ALC-related patients for HCC diagnosis since DCP has been found to have a poorer AUC value in nonviral infected patients.⁵

Differentially Expressed CF Peptides in HCV-Related HCC Compared to HCV-Related Cirrhosis

In HCV-related samples 12 CF peptides from nine CF proteins were found to be upregulated in HCC compared to cirrhosis samples, whereas 14 CF peptides from 13 CF proteins were downregulated (Table 3). The top network of these changed proteins involves 11 proteins: cadherin-5 (CDH5), P-selectin (SELP), heparin cofactor 2 (SERPIND1), antithrombin-III (SERPINC1), plasma protease C1 inhibitor (SERPING1), fibromodulin (FMOD), receptor type-protein tyrosine phosphatase S (PTPRS), tenascin C (TNC), collagen alpha-1(VI) chain (COL6A1), alpha-2-macroglobulin (A2M), and growth/differentiation factor 8 (MSTN). These proteins are involved interactively in tissue development, cardiovascular system development and function, and cancer (Figure 3B).

Among the 22 differentially expressed CF proteins, CF hemopexin was found to be upregulated and has been considered to be a possible biomarker for liver disease.^{30,31} In our study, neither of the two CF peptides from hemopexin provided promising AUC values to distinguish HCC from cirrhosis patients, where the two CF peptides showed AUC values of 0.70 at site 187 (CF peptide: sWPAVGncSSALR) and 0.74 at site 453 (CF peptide: aLPQPQnVTSLLGcTH) (Figure S6). The CF level of cadherin-5 (CD144) at site 61 (CF peptide: dWIWNQmHIDEEnTSLPHHVGk) provides the highest AUC value of 0.75 with a specificity of 75.0% at a sensitivity of 83.3% to distinguish HCV-related HCC from HCV-

related cirrhosis (Figure S7). A combined analysis with clinical AFP value analysis showed an AUC value of 0.80 (LR test $p = 0.02$), whereas AFP alone in these samples provides an AUC value of 0.64. Cadherin-5 is a cell adhesion protein and functions in connecting cells of heterogeneous cell types.³² Glycosylated cadherin-5 has been identified as a novel biomarker of metastatic breast cancer.³³ CF cadherin-5 in this study showed some diagnostic value for HCV-related HCC samples. However, our strategy to screen for CF peptides as biomarkers for HCV-related samples did not discover potential biomarkers that are as promising as that for ALC-related samples. Some other differentially expressed CF peptides in ALC- or HCV-related HCC samples compared with cirrhosis samples are shown in Figure S8.

Analysis of the Three CF Sites of Fibronectin in Pooled ALC-Related and HCV-Related Samples

The expression level of three CF sites of fibronectin in pooled ALC-related and HCV-related samples is shown in Figure 6. The average CF levels of all the three sites were slightly higher in HCC compared with cirrhosis samples, but there were no significant differences, whereas with etiology-separated analysis, the CF level of fibronectin at site 1007 combined with the AFP value provides a potentially significant marker for early HCC diagnosis in ALC-related cirrhosis patients. It is thus essential to separate HCC samples with different etiologies to avoid masking potential changes that could serve as markers.

CONCLUSIONS

A large-scale screening for site-specific CF of serum proteins as markers for early stage HCC was performed in patients of either ALC-related or HCV related etiologies. In sum, 1300 CF peptides from 613 CF proteins were identified, where 20 CF peptides were differentially expressed in ALC-related HCC samples compared with ALC-related cirrhosis samples and 26 CF peptides changed in HCV-related HCC compared with HCV-related cirrhosis samples. Among these, three CF peptides from fibronectin were upregulated in ALC-related HCC compared to ALC-related cirrhosis with an AUC value of 0.89 at site 1007. When combined with the AFP index, the AUC value reached to 0.92 at site 1007, much improved compared to that using AFP alone (LR test $p < 0.001$). The CF peptides of fibronectin may serve as potential biomarkers for early stage HCC screening in ALC-related cirrhosis patients. This change was not found in HCV-related samples. In HCV-related samples, the CF level of cadherin-5 at site 61 provides the best result. A combined analysis with AFP shows an AUC value of 0.80 to distinguish HCV-related HCC from HCV-related cirrhosis, which is improved compared to that with AFP alone (LR test $p = 0.02$). Further validation of these potential markers needs to be performed using larger sample sets.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

This work was funded by the National Cancer Institute under grant nos. 1R01 CA160254 (D.M.L.) and 1R01 CA154455 (D.M.L.) and received partial support from the National Institutes of Health through grant no. R01 GM

49500 (D.M.L.). The data was generated by the Proteomics Core at Wayne State University. The Core is supported by Center grant nos. P30 ES020957 and P30 CA022453 and Instrumentation grant no. S10 010700.

ABBREVIATIONS

HCC	hepatocellular carcinoma
ALC	alcohol
HCV	hepatitis C virus
AFP	alpha-fetoprotein
LCA	<i>Lens culinaris</i> Agglutinin
AFP-L3	LCA-bound fraction of AFP
LC-MS/MS	liquid chromatography with tandem mass spectrometry
HCD	high-energy collision dissociation
ROC	receiver operating characteristic
AUC	area under the curve
TCEP	tris(2-carboxyethyl)phosphine
IAA	iodoacetamide
TEAB	triethylammonium bicarbonate
TFA	trifluoroacetic acid
LR test	likelihood ratio test

REFERENCES

1. Kamangar F, Dores GM, Anderson WF. Patterns of cancer incidence, mortality, and prevalence across five continents: Defining priorities to reduce cancer disparities in different geographic regions of the world. *J. Clin. Oncol.* 2006; 24:2137–2150. [PubMed: 16682732]
2. El-Serag HB. CURRENT CONCEPTS Hepatocellular Carcinoma. *N. Engl. J. Med.* 2011; 365:1118–1127. [PubMed: 21992124]
3. Llovet JM, Burroughs A, Bruix J. Hepatocellular carcinoma. *Lancet.* 2003; 362:1907–1917. [PubMed: 14667750]
4. Trevisani F, D'Intino PE, Morselli-Labate AM, Mazzella G, Accogli E, Caraceni P, Domenicali M, De Notariis S, Roda E, Bernardi M. Serum alpha-fetoprotein for diagnosis of hepatocellular carcinoma in patients with chronic liver disease: influence of HBsAg and anti-HCV status. *J. Hepatol.* 2001; 34:570–575. [PubMed: 11394657]
5. Marrero JA, Feng ZD, Wang YH, Nguyen MH, Befeler AS, Roberts LR, Reddy KR, Harnois D, Llovet JM, Normolle D, Dalhgren J, Chia D, Lok AS, Wagner PD, Srivastava S, Schwartz M. alpha-Fetoprotein, Des-gamma Carboxyprothrombin, and Lectin-Bound alpha-Fetoprotein in Early Hepatocellular Carcinoma. *Gastroenterology.* 2009; 137:110–118. [PubMed: 19362088]
6. Wu J, Zhu JH, Yin HD, Buckanovich RJ, Lubman DM. Analysis of Glycan Variation on Glycoproteins from Serum by the Reverse Lectin-Based ELISA Assay. *J. Proteome Res.* 2014; 13:2197–2204. [PubMed: 24575722]
7. Li QK, Gabrielson E, Zhang H. Application of glycoproteomics for the discovery of biomarkers in lung cancer. *Proteomics: Clin. Appl.* 2012; 6:244–256. [PubMed: 22641610]

8. Pompach P, Brnakova Z, Sanda M, Wu J, Edwards N, Goldman R. Site-specific Glycoforms of Haptoglobin in Liver Cirrhosis and Hepatocellular Carcinoma. *Mol. Cell. Proteomics*. 2013; 12:1281–1293. [PubMed: 23389049]
9. Lin ZX, Yin HD, Lo A, Ruffin MT, Anderson MA, Simeone DM, Lubman DM. Label-free relative quantification of alpha-2-macroglobulin site-specific core-fucosylation in pancreatic cancer by LC-MS/MS. *Electrophoresis*. 2014; 35:2108–2115. [PubMed: 24285556]
10. Tan ZJ, Yin HD, Nie S, Lin ZX, Zhu JH, Ruffin MT, Anderson MA, Simeone DM, Lubman DM. Large-Scale Identification of Core-Fucosylated Glycopeptide Sites in Pancreatic Cancer Serum Using Mass Spectrometry. *J. Proteome Res*. 2015; 14:1968–1978. [PubMed: 25732060]
11. Nie S, Lo A, Wu J, Zhu JH, Tan ZJ, Simeone DM, Anderson MA, Shedden KA, Ruffin MT, Lubman DM. Glycoprotein Biomarker Panel for Pancreatic Cancer Discovered by Quantitative Proteomics Analysis. *J. Proteome Res*. 2014; 13:1873–1884. [PubMed: 24571389]
12. Ahn JM, Sung HJ, Yoon YH, Kim BG, Yang WS, Lee C, Park HM, Kim BJ, Kim BG, Lee SY, An HJ, Cho JY. Integrated Glycoproteomics Demonstrates Fucosylated Serum Paraoxonase 1 Alterations in Small Cell Lung Cancer. *Mol. Cell. Proteomics*. 2014; 13:30–48. [PubMed: 24085812]
13. Lee HJ, Cha HJ, Lim JS, Lee SH, Song SY, Kim H, Hancock WS, Yoo JS, Paik YK. Abundance-Ratio-Based Semiquantitative Analysis of Site-Specific N-Linked Glycopeptides Present in the Plasma of Hepatocellular Carcinoma Patients. *J. Proteome Res*. 2014; 13:2328–2338. [PubMed: 24628331]
14. Zhu JH, Lin ZX, Wu J, Yin HD, Dai JL, Feng ZD, Marrero J, Lubman DM. Analysis of Serum Haptoglobin Fucosylation in Hepatocellular Carcinoma and Liver Cirrhosis of Different Etiologies. *J. Proteome Res*. 2014; 13:2986–2997. [PubMed: 24807840]
15. Nakagawa T, Miyoshi E, Yakushijin T, Hiramatsu N, Igura T, Hayashi N, Taniguchi N, Kondo A. Glycomic analysis of alpha-fetoprotein L3 in hepatoma cell lines and hepatocellular carcinoma patients. *J. Proteome Res*. 2008; 7:2222–2233. [PubMed: 18479159]
16. Toyoda H, Kumada T, Tada T, Kaneoka Y, Maeda A, Kanke F, Satomura S. Clinical utility of highly sensitive Lens culinaris agglutinin-reactive alpha-fetoprotein in hepatocellular carcinoma patients with alpha-fetoprotein < 20 ng/mL. *Cancer Sci*. 2011; 102:1025–1031. [PubMed: 21244578]
17. Liu YS, He JT, Li C, Benitez R, Fu S, Marrero J, Lubman DM. Identification and Confirmation of Biomarkers Using an Integrated Platform for Quantitative Analysis of Glycoproteins and Their Glycosylations. *J. Proteome Res*. 2010; 9:798–805. [PubMed: 19961239]
18. Yin HD, Lin ZX, Nie S, Wu J, Tan ZJ, Zhu JH, Dai JL, Feng ZD, Marrero J, Lubman DM. Mass-Selected Site-Specific Core-Fucosylation of Ceruloplasmin in Alcohol-Related Hepatocellular Carcinoma. *J. Proteome Res*. 2014; 13:2887–2896. [PubMed: 24799124]
19. Fanayan S, Hincapie M, Hancock WS. Using lectins to harvest the plasma/serum glycoproteome. *Electrophoresis*. 2012; 33:1746–1754. [PubMed: 22740463]
20. Morgan TR, Mandayam S, Jamal MM. Alcohol and hepatocellular carcinoma. *Gastroenterology*. 2004; 127:S87–S96. [PubMed: 15508108]
21. Riener MO, Stenner F, Liewen H, Soll C, Breitenstein S, Pestalozzi BC, Samaras P, Probst-Hensch N, Hellerbrand C, Mullhaupt B, Clavien PA, Bahra M, Neuhaus P, Wild P, Fritzsche F, Moch H, Jochum W, Kristiansen G. Golgi Phosphoprotein 2 (GOLPH2) Expression in Liver Tumors and Its Value as a Serum Marker in Hepatocellular Carcinomas. *Hepatology*. 2009; 49:1602–1609. [PubMed: 19291786]
22. Mao YL, Yang HY, Xu HF, Lu X, Sang XT, Du SD, Zhao HT, Chen W, Xu Y, Chi TY, Yang ZY, Cai JQ, Li H, Chen JG, Zhong SX, Mohanti SR, Lopez-Soler R, Millis JM, Huang JF, Zhang HB. Golgi protein 73 (GOLPH2) is a valuable serum marker for hepatocellular carcinoma. *Gut*. 2010; 59:1687–1693. [PubMed: 20876776]
23. Marrero JA, Romano PR, Nikolaeva O, Steel L, Mehta A, Fimmel CJ, Comunale MA, D'Amelio A, Lok AS, Block TM. GP73, a resident Golgi glycoprotein, is a novel serum marker for hepatocellular carcinoma. *J. Hepatol*. 2005; 43:1007–1012. [PubMed: 16137783]

24. Drake RR, Schwegler EE, Malik G, Diaz J, Block T, Mehta A, Semmes OJ. Lectin capture strategies combined with mass spectrometry for the discovery of serum glycoprotein biomarkers. *Mol. Cell. Proteomics*. 2006; 5:1957–1967. [PubMed: 16760258]
25. Iftikhar R, Kladney RD, Havlioglu N, Schmitt-Graff A, Gusmirovic I, Solomon H, Luxon BA, Bacon BR, Fimmel CJ. Disease- and cell-specific expression of GP73 in human liver disease. *Am. J. Gastroenterol.* 2004; 99:1087–1095. [PubMed: 15180730]
26. Marrero JA, Su GL, Wei W, Emick D, Conjeevaram HS, Fontana RJ, Lok AS. Des-gamma carboxyprothrombin can differentiate hepatocellular carcinoma from nonmalignant chronic liver disease in American patients. *Hepatology*. 2003; 37:1114–1121. [PubMed: 12717392]
27. To WS, Midwood KS. Plasma and cellular fibronectin: distinct and independent functions during tissue repair. *Fibrog. Tissue Repair*. 2011; 4:21.
28. Varki, A.; Cummings, RD.; Esko, JD.; Freeze, HH.; Stanley, P.; Bertozzi, CR.; Hart, GW.; Etzler, ME., editors. *Essentials of Glycobiology*. New York: Cold Spring Harbor Laboratory Press; 2008. p. 107
29. Noda K, Miyoshi E, Uozumi N, Yanagidani S, Ikeda Y, Gao CX, Suzuki K, Yoshihara H, Yoshikawa M, Kawano K, Hayashi N, Hori M, Taniguchi N. Gene expression of alpha 1–6 fucosyltransferase in human hepatoma tissues: A possible implication for increased fucosylation of alpha-fetoprotein. *Hepatology*. 1998; 28:944–952. [PubMed: 9755230]
30. Kobayashi S, Nouse K, Kinugasa H, Takeuchi Y, Tomoda T, Miyahara K, Hagihara H, Kuwaki K, Onishi H, Nakamura S, Ikeda F, Miyake Y, Shiraha H, Takaki A, Yamamoto K. Clinical utility of serum fucosylated hemopexin in Japanese patients with hepatocellular carcinoma. *Hepatol. Res.* 2012; 42:1187–1195. [PubMed: 22639859]
31. Benicky J, Sanda M, Pompach P, Wu J, Goldman R. Quantification of Fucosylated Hemopexin and Complement Factor H in Plasma of Patients with Liver Disease. *Anal. Chem.* 2014; 86:10716–10723. [PubMed: 25302577]
32. Lampugnani MG, Orsenigo F, Rudini N, Maddaluno L, Boulday G, Chapon F, Dejana E. CCM1 regulates vascular-lumen organization by inducing endothelial polarity. *J. Cell Sci.* 2010; 123:1073–1080. [PubMed: 20332120]
33. Fry SA, Sinclair J, Timms JF, Leatham AJ, Dwek MV. A targeted glycoproteomic approach identifies cadherin-5 as a novel biomarker of metastatic breast cancer. *Cancer Lett.* 2013; 328:335–344. [PubMed: 23079531]

Figure 1. Number of peptides identified (in percentage) as a function of the number of runs using Q Exactive MS. This figure illustrates the need to run a sample at least three times. (A) Sample run three times and (B) sample run for four times.

Figure 2.

Number of identified CF peptides (in percentage) as a function of the number of sample sets used in this study. This figure illustrates that nine sets of samples are sufficient to identify more than 95% of the total identified CF peptides in either ALC-related or HCV-related samples.

Figure 4.

Analysis of the CF level of fibronectin at sites 528 (A), 542 (B), and 1007 (C). (top) The CF level of fibronectin at three sites in ALC-related HCC and ALC-related cirrhosis. A *t*-test comparison was made between the two groups. Error bar indicates SEM; ***, $p < 0.001$. (bottom) ROC analysis of the CF level of fibronectin at three sites in ALC-related HCC compared with ALC-related cirrhosis.

Figure 5.

Combined analysis of the CF level of fibronectin at sites 528 (A), 542 (B), and 1007 (C) and clinical AFP level. (top) Combined ROC analysis of the two indexes in ALC-related HCC compared with ALC-related cirrhosis. (bottom) Two-dimensional plot of the two indexes in ALC-related HCC compared with ALC-related cirrhosis.

Figure 6.

CF level of fibronectin at sites 528, 542, and 1007 in pooled ALC-related and HCV-related HCC or cirrhosis. A *t*-test comparison was made between the two groups. Error bar indicates SD. No significant differences were found.

Table 1

Clinical Characteristics of the 52 Patients Involved in This Study

disease group	ALC-related ^c		HCV-related ^d	
	cirrhosis	HCC	cirrhosis	HCC
sample size ^a	14	14	12	12
age (median, range)	64 (57–70)	63 (48–71)	57 (48–74)	61 (49–80)
gender F/M	4:10	5:9	4:8	4:8
AFP ^b (median, ng/mL)	2.45	5.95	4.10	6.40
MELD score	9.4 ± 3.2	10.4 ± 3.5	10.0 ± 2.7	8.1 ± 3.3
BCLC stage	NA	NA	A	A

^aSamples were provided by the Division of Gastroenterology, University of Michigan.^bAFP level was provided by the Division of Gastroenterology.^cALC, alcohol.^dHCV, hepatitis C virus.

Table 2
Differentially Expressed CF Peptides in ALC-Related HCC Compared with ALC-Related Cirrhosis

accession	protein name	CF peptides ^a	time ^b	p value ^c	fold changed ^d
P02751	Fibronectin	hEEGHmLncTcFGQGR	14	0.000	1.96
		IDAPTNLQFVnETDSTVLVR	14	0.000	2.09
		dQcIVDDITYNVnDTFHK	14	0.000	1.81
Q8N6C8	Leukocyte immunoglobulin-like receptor subfamily A member 3 (CD85e)	qPQAGLSQAnFTLGPVSR	13	0.008	1.38
P10909	Clusterin(Apo-J) (Testosterone-repressed prostate message 2)	eLPGVcnETmmALWEEckPeLk	14	0.011	1.38
		mLnTSSLLLEQLNEQFNWVSR	6	0.041	1.76
		kEDALnETR	9	0.042	1.27
		lAnLTQGEDQYYLR	12	0.043	1.25
P05543	Thyroxine-binding globulin	vTAcHSSQPnATLYk	14	0.021	1.45
P32942	Intercellular adhesion molecule 3 (CD50)	tELDmQPQGLGLFVnTSAPR	6	0.041	1.39
Q12860	Contactin-1 (Neural cell surface protein F3)	gTEWLVnSSR	14	0.028	1.16
Q8NBJ4	Golgi membrane protein 1 (Golgi membrane protein GP73)	aVLVNnITTGER	11	0.029	1.31
Q15223	Nectin-1	NPhGTVTVISR	6	0.026	1.33
Q14314	Fibroleukin	lHVGNYNnGTAGDALR	7	0.018	1.31
P04275	von Willebrand factor	aSPSSScnISSGEmQk	14	0.039	1.26
P17936	Insulin-like growth factor-binding protein 3	ykVDYESQSTDTQnFSSES	6	0.026	0.33
P09172	Dopamine beta-hydroxylase	sLEAlnGSLQmGLQR	14	0.009	0.50
P00451	Coagulation factor VIII	kGEEENLEGLGnQTK	6	0.026	0.73
P07996	Thrombospondin-1	vScPImpcSnATVPDGEccPR	14	0.014	0.75
P00734	Prothrombin	yPHkPEInSTTHPGADLQENFcR	8	0.038	0.82

^a Amino acids with modification are indicated with lower case. N-terminal is iTRAQ labeled; m indicates methionine oxidation; n indicates core-fucosylated asparagine; c indicates carbamidomethylation of cysteine.

^b Time was detected in the 14 sample sets of ALC-related samples.

^c p values were calculated as described in statistical analysis in Materials and Methods.

^d Fold change was calculated by the average value of HCC samples divided by the average value of cirrhosis samples.

Table 3
Differentially Expressed CF Peptides in HCV-Related HCC Compared with HCV-Related Cirrhosis

accession	protein name	CF peptides ^a	time ^b	p value ^c	fold changed ^d
P02790	Hemopexin	aLPQqQnVTSLLGcTH	12	0.015	1.65
		sWPAVGncSSALR	12	0.044	1.59
P03952	Plasma kallikrein (Kininogenin)	gVNFnVSk	11	0.021	1.64
P05155	Plasma protease C1 inhibitor	dTFVnASR	12	0.017	1.53
		vLSnSSDANLELINTWVAK	12	0.031	1.42
P05546	Heparin cofactor 2	gGETAQsADPQWEQLNNknLSmPLLPADFHk	12	0.014	1.53
		dFVnASSk	12	0.036	1.74
P33151	Cadherin-5 (CD144)	dWIWNQmHIDEEknTSLPHHVgk	12	0.006	1.22
P43652	Afamin (Alpha-albumin) (Alpha-Alb)	dIENFnSTQk	12	0.045	2.02
P01008	Antithrombin-III (ATIII) (Serpin C1)	aAINkWWVSnkTEGR	10	0.035	1.55
P16109	P-selectin (CD62 antigen-like family member P)	gnmTcLHSAk	6	0.015	1.46
P01023	Alpha-2-macroglobulin (Alpha-2-M)	gNEANYYSnATTDEHGLVQFSInTTNVmGTSLTVR	12	0.045	1.23
P41271	Neuroblastoma suppressor of tumorigenicity 1	nITQIVGHSGcEAK	12	0.028	0.75
O14498	Immunoglobulin superfamily containing leucine-rich repeat protein	fQAFArGSLIPDFGk	10	0.016	0.86
Q92823	Neuronal cell adhesion molecule	qkDGDDEWTSVVVAnVSk	11	0.018	0.78
Q6UXB8	Peptidase inhibitor 16	sLPNFPnTSATAnATGGR	12	0.020	0.68
		eHYnLSAATcSPGQmcGHYTTQVVWAK	10	0.039	0.73
O00451	GDNF family receptor alpha-2	nAIQAFGnGTDVNVSPk	11	0.026	0.83
Q13332	Receptor-type tyrosine-protein phosphatase S	kVEAEALnATAIR	11	0.026	0.77
P07998	Ribonuclease pancreatic (RNase A)	snSSmHITDcR	12	0.024	0.70
P24821	Tenascin	vEAAQnLTLPGSLR	12	0.033	0.84
Q06828	Fibromodulin	IYLDHnNlTR	11	0.040	0.82
O14793	Growth/differentiation factor 8 (Myostatin)	IETAPnISk	11	0.042	0.73
Q14314	Fibroleukin	vAnLTFVVNSLDGk	9	0.045	0.70
P12109	Collagen alpha-1 (VI) chain	nFTAADWQQR	8	0.028	0.61
Q12841	Follistatin-related protein 1 (Follistatin-like protein 1)	fVEQnETAnITTYPDQENnk	7	0.038	0.81

^a Amino acids with modification are indicated with lower case. N-terminal is ITRAQ labeled; m indicates methionine oxidation; n indicates core-fucosylated asparagine; c indicates carbamidomethylation of cysteine.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

t_p Time was detected in the 12 sample sets of HCV-related samples.
 p_c p values were calculated as described in statistical analysis in Materials and Methods.
 d_f Fold change was calculated by the average value of HCC samples divided by the average value of cirrhosis samples.