

Subcellular Tissue Proteomics of Hepatocellular Carcinoma for Molecular Signature Discovery

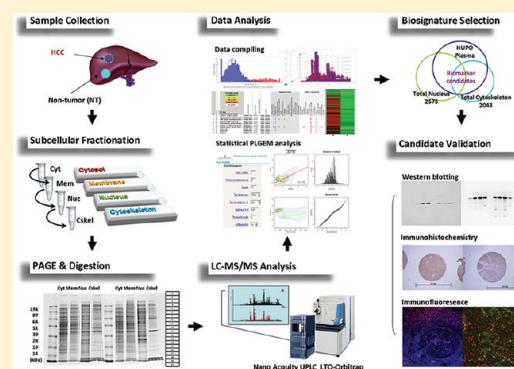
Yong-Yook Lee,^{†,‡} Kimberly Q. McKinney,^{†,‡} Sriparna Ghosh,^{§,‡} David A. Iannitti,^{§,||} John B. Martinie,^{§,||} F. Ryan Caballes,^{§,†} Mark W. Russo,^{§,†} William A. Ahrens,[#] Deborah H. Lundgren,[¶] David K. Han,[¶] Herbert L. Bonkovsky,^{§,‡,†} and Sun-Il Hwang^{*,†,§,‡}

[†]Proteomics Laboratory for Clinical and Translational Research, [§]Liver-Biliary-Pancreatic Center, [‡]Departments of Research, ^{||}General Surgery, [¶]Internal Medicine, and [#]Pathology, Carolinas HealthCare System, Charlotte, NC,

^{*}Department of Cell Biology and Center for Vascular Biology, University of Connecticut Health Center, Farmington, CT

Supporting Information

ABSTRACT: Hepatocellular carcinoma (HCC) is one of the leading causes of mortality from solid organ malignancy worldwide. Because of the complexity of proteins within liver cells and tissues, the discovery of therapeutic targets of HCC has been difficult. To investigate strategies for decreasing the complexity of tissue samples for detecting meaningful protein mediators of HCC, we employed subcellular fractionation combined with 1D-gel electrophoresis and liquid chromatography–tandem mass spectrometry analysis. Moreover, we utilized a statistical method, namely, the Power Law Global Error Model (PLGEM), to distinguish differentially expressed proteins in a duplicate proteomic data set. Mass spectrometric analysis identified 3045 proteins in nontumor and HCC from cytosolic, membrane, nuclear, and cytoskeletal fractions. The final lists of highly differentiated proteins from the targeted fractions were searched for potentially translocated proteins in HCC from soluble compartments to the nuclear or cytoskeletal compartments. This analysis refined our targets of interest to include 21 potential targets of HCC from these fractions. Furthermore, we validated the potential molecular targets of HCC, MATR3, LETM1, ILF2, and IQGAP2 by Western blotting, immunohistochemistry, and immunofluorescent microscopy. Here we demonstrate an efficient strategy of subcellular tissue proteomics toward molecular target discovery of one of the most complicated human disease, HCC.



KEYWORDS: hepatocellular carcinoma, molecular signature, tissue proteomics, subcellular fractionation, protein localization and translocation, PLGEM

INTRODUCTION

Hepatocellular carcinoma (HCC) is one of the leading causes of mortality from solid organ malignancy worldwide.^{1–3} There are well-recognized environmental risk factors such as chronic hepatitis B and C viral infections, exposure to aflatoxin, and other etiologic factors for HCC. Recognized host factors include a myriad of genetic or acquired metabolic disorders that lead to hepatic fibrosis or cirrhosis.⁴ The diagnosis of primary HCC is occurring with increasing frequency in the United States.^{2,5} The five year survival rate of HCC remains less than 10%. The poor survival rate associated with HCC is primarily related to the lack of reliable molecular targets for treatment and early disease markers.^{6,7} Early detection of HCC may improve the five year survival rate.⁸ However, detection during the early stages of HCC is difficult because of lack of specific symptoms in early disease and the limited information of the disease heterogeneity and etiology.⁹ Serum alpha-feto-protein (AFP) is the most widely used biosignature for HCC, but it is far from perfect, being elevated in less than 60% of patients with HCC and also

sometimes elevated in those without HCC.¹⁰ These challenges have prompted investigators to pursue studies aimed at identifying additional diagnostic and therapeutic targets.

One of the strategies for biosignature discovery in HCC is analysis of diseased and control tissues by semiquantitative proteomic approaches. In general, global profiling and semi-quantification using liquid chromatography–tandem mass spectrometry (LC–MS/MS) can identify potential molecular targets differentially expressed in cancers. Proteomic analyses of HCC tissues have been performed using traditional 2D-PAGE followed by MS analysis. In one recent study, this approach led to the identification of a number of proteins modified in HCC.⁹ Specifically, 1068 proteins were identified in 80 pairs of HCC tumor and matched nontumor tissues.¹¹ Additional studies have explored the proteome-wide patterns of HCC to identify biosignatures, one of which showed 200 common proteins in

Received: May 31, 2011

Published: September 13, 2011

Table 1. Selected Demographic, Histologic, and Biochemical Features of Subjects Studied

subject #	23	55	65
Gender	Male	Male	Female
Age (y)	60	76	66
HCC ^a Stage	ypT1 N0MX	pT3 NX MX	T3 NX MX
Hepatitis C	Unknown	None	None
Surgical Procedure	Left hepatic lobectomy	Left hepatic lobectomy	Left hepatic lobectomy with portal vein resection/reconstruction
Underlying Liver Disease	Mild nonspecific hepatitis	NAFLD and heterozygous C282Y mutation of HFE	None
Histologic Characteristics of HCC	Moderately differentiated	Moderately to poorly differentiated	Poorly differentiated
Histopathologic Characteristics of NT ^b	The liver uninvolved by tumor is essentially unremarkable.	The uninvolved hepatic parenchyma is brown and slightly granular.	Adjacent parenchyma is red-brown and unremarkable
Blood Chemistries			
ALT ^c (IU/L)	124	53	124
AST ^d (IU/L)	80	63	80
AP ^e (IU/L)	126	134	126
Total Bilirubin(mg/dL)	1	1	1
Albumin (g/dL)	4	4	4
Total Protein(g/dL)	7	8	7
INR ^f	1	1	1

^a HCC: Hepatocellular carcinoma ^b NT: Nontumor liver. ^c ALT: Alanine aminotransferase. ^d AST: Aspartate aminotransferase. ^e AP: Alkaline phosphatase. ^f INR: International normalized ratio. All subjects were Caucasian.

fractionated liver samples, 639 proteins from tissue array samples and 1424 proteins employing stable isotope labeling with amino acids in cell culture (SILAC) in 3 human liver cancer cell lines.^{12–14} Other studies showed confident identification of 1800–2210 proteins from subfractions of mouse liver tissue.^{15,16} These reports notwithstanding, protein profiling has not fully overcome the challenge of the large dynamic range in achieving the identification of low-abundance proteins expressed in mammalian tissue. The complexity of the peptide mixtures often overwhelms the mass spectrometer and the analytical capacity of the software, and low-abundance peptides may be masked when mixed with peptides of high abundance.^{17,18} Contemporary cancer proteomics is aimed at identifying changes in protein expression, structure, and post-translational modification, as well as subcellular localization.⁷ Recently established, the “Liverbase” (<http://liverbase.hupo.org.cn>) integrates information on the human liver proteome, including function, abundance, and associated disease information.¹⁹ Other groups showed that fractionation of cell lines is a powerful tool for protein localization and translocation, which facilitates answering specific biological questions.^{20,21} These findings prompted us to investigate strategies to decrease the complexity of tissue proteomes in cancer in order to generate meaningful identification and semiquantification of the proteins; moving toward the goal of establishing molecular profiling of the disease.²² In this study, we employed subcellular fractionation and LC tandem mass spectrometry (MS/MS) analysis to identify specific proteins in HCC tissues, utilizing protein localization and putative translocation information as a novel denominator of biosignature discovery. Moreover, we utilized the statistical method, Power Law Global Error Model (PLGEM), for spectral count quantification which provided not only an efficient and informative screening strategy, but also reliable and robust biosignature mapping of HCC.^{23–25}

■ EXPERIMENTAL PROCEDURES

Reagents and Chemicals

All ultra performance liquid chromatography (UPLC) solvents, 0.1% formic acid in water and 0.1% in ACN were of the LC–MS grade (OmniSolv) purchased from EMD (Gibbstown, NJ) and used without further purification. Modified trypsin (Sequencing grade) was from Promega (Madison, WI) and GelCode blue stain reagent was from Pierce (Rockford, IL). Complete protease inhibitor cocktail tablet was obtained from Roche (Mannheim, Germany). BSA protein assay kit and Restore plus Western blot stripping buffer were purchased from Thermo Scientific (Rockford, IL). FractionPREP cell fractionation kit was from Biovision (Mountain View, CA) and PBS, Alexa 594 and prolong gold antifade 4',6-diamidino-2-phenylindole were from Invitrogen (Carlsbad, CA).

Ammonium bicarbonate, ammonium acetate, DTT, iodoacetamide, Tris-HCl, bromophenol blue, β-mercaptoethanol, Tween 20, 3% H₂O₂ and SDS were obtained from Sigma-Aldrich (St. Louis, MO). Glycerol was from Life Technologies (Gaithersburg, MD) and ECL Advance kit was from GE Healthcare (Buchinghamshire, U.K.). Primary antibodies against human LETM1 and EF-hand domain-containing protein 1, mitochondrial, LETM1 (mouse monoclonal); Thyroid hormone receptor-associated protein 3, TRAP150 (goat polyclonal); Carbamoyl-phosphate synthase [ammonia], mitochondrial, CPS1 (goat polyclonal); Fatty acid synthase, FAS (rabbit polyclonal); Ras GTPase-activating-like protein IQGAP2, IQGAP2 (mouse monoclonal); Matrin3, MATR3 (goat polyclonal); Pan-actin (goat polyclonal); Glyceraldehyde 3-phosphate dehydrogenase and secondary antibodies were purchased from Santa Cruz (Santa Cruz, CA). Antibodies against human Periostin, OSF2 (rabbit polyclonal); Nuclear mitotic apparatus protein 1, NUMA (mouse monoclonal) and

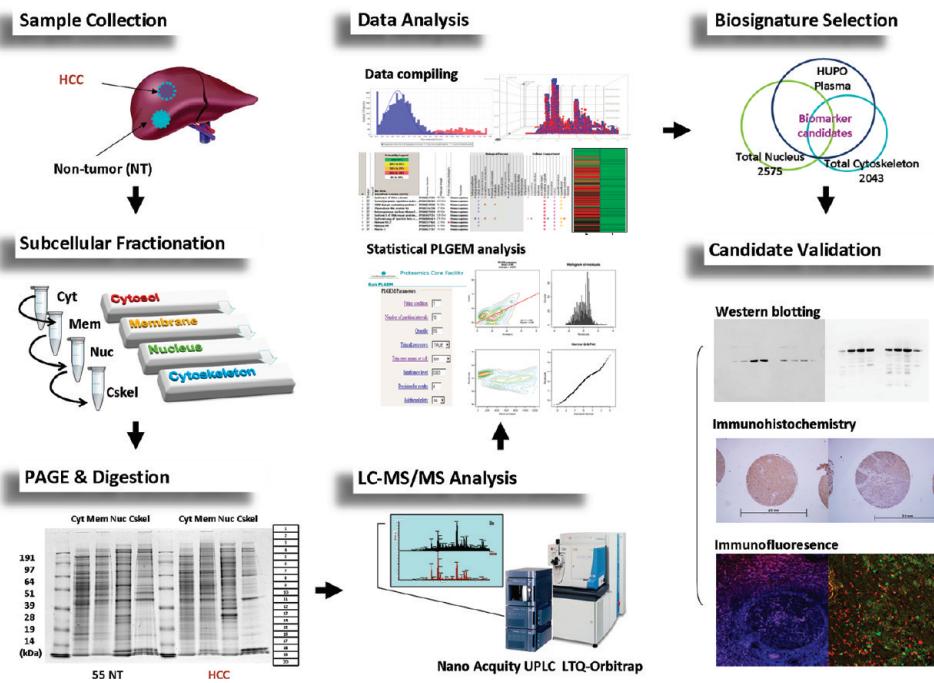


Figure 1. Schematic diagram: procedures and Methods used for subcellular proteomics for HCC biosignature discovery.

Interleukin enhancer-binding factor 2, ILF2 (rabbit polyclonal) were obtained from Abcam (Cambridge, MA). AutostainerPlus, peroxidase-conjugated streptavidin and 3'-diaminobenzidine were from Dako (Carpinteria, CA). All buffers and solutions were prepared using deionized water by Milli-Q, Millipore (Bedford, MA).

Human Liver Tissue Collection and Cell Culture

Patients presenting to Carolinas Medical Center (Charlotte, NC) for treatment of HCC were recruited and consented to participate in this study based on Institutional Review Board (IRB) approved protocols. There were no exclusion criteria for age, race, or gender. However, patients under age 18 years required parental consent. When participants underwent surgery for resection, the resected liver was sent to the pathology department for staging. At this time the pathologist dissected a part of the HCC tumor as well as a part of adjacent nontumor (NT) liver tissue from each patient and snap froze them in liquid nitrogen. All samples collected were inventoried in the Liver-Biliary-Pancreatic tissue repository located at Cannon Research Center of Carolinas Medical Center. Tissues from 3 patients without appreciable underlying liver disease (numbers 23, 55, and 65) were used for this study. Each patient's medical record was reviewed and the salient demographic and clinical pathology for each patient was recorded (Table 1).

Human hepatoma cell lines, Huh7.5 and CON1, used for proteomic sequencing by subcellular fractionation, and Huh-7, used for immunofluorescent staining, were cultured following Hou et al.²⁶

Subcellular Protein Fractionation

An overview of the methods and procedures employed is shown in Figure 1. While on dry ice, approximately 20 mg of tissue was excised from snap-frozen nontumor (NT) and HCC liver tissues from the three patients (hereafter referred to by study ID numbers as 23-NT, 23-HCC, 55-NT, 55-HCC and 65-NT, 65-HCC). The FractionPREP cell fractionation kit was used

for subcellular fractionation based on the manufacturer's protocol. Briefly, each tissue sample was minced using a scalpel and the pieces were transferred to a 2.0 mL tube then washed with 1× ice-cold PBS 2 times for 2 min on ice. Before starting the procedure, 2 μL protease inhibitor cocktail and 2 μL DTT were added to 1 mL of cytosol extraction buffer (CEB), membrane extraction buffer (MEB) A and nuclear extraction buffer (NEB). Samples were homogenized using a TH115 homogenizer (Omni International, Kennesaw, GA) for 30 s to 1 min until tissue was completely dissociated in the 400 μL CEB. The samples were then incubated on ice 20 min with gentle tapping 3–4 times every 5 min after pipetting several times to mix well. The homogenate was centrifuged at 700× g for 10 min at 4 °C (Spin 1). The supernatant was transferred to a clean tube and labeled Fraction "Cyt" (cytosol fraction). To the pellet was added 400 μL MEB, mixed after adding 22 μL of MEB B and then incubated on ice for 1 min. The solution was centrifuged at 1000× g at 4 °C for 5 min (Spin 2). The supernatant was transferred to a clean tube and labeled fraction "Mem" (membrane/particulate fraction). To the pellet was added 200 μL NEB. The pellet was resuspended and kept on ice for 40 min with constant shaking. The samples were then centrifuged at 18 000× g for 10 min at 4 °C (Spin 3). The supernatant was transferred to a clean tube and labeled fraction "Nuc" (nuclear fraction). To the pellet was added 100 μL of 0.2% SDS containing 10 mM DTT (Sup. Figure 1A, Supporting Information). The pellet was resuspended and this sample labeled fraction "Cske" (cytoskeleton fraction). All fractions were stored at –80 °C for future use. Protein quantitation was carried out using a microplate procedure and the bicinchoninic acid assay, with BSA as the protein standard, using a 1:10 dilution of each fraction. The cell line fractionation was carried out according to the same method as described above.

1-D SDS PAGE and In-gel Digestion

The subcellular fractions from one cancerous (HCC) and nontumor (NT) tissue pair (55-NT/-HCC) were chosen for

representative LC–MS/MS analysis. From the four subcellular fractions of each 55-NT and 55-HCC samples, total proteins (30 µg) were separated on 10% Bis-Tris NuPAGE gels (Invitrogen, Carlsbad, CA) with 6 X sample buffer containing 300 mM Tris-HCl, 0.01% bromophenol blue (w/v), 15% glycerol (v/v), 6% SDS (w/v) and 1% beta mercaptoethanol (v/v) after boiling for 10 min at 95 °C. The gel was run at 150 V for 60 min. Gels were washed with deionized water, and then fixed in 50% methanol, 7% acetic acid for 15 min. After fixation, gels were rinsed with deionized water then stained using the Gelcode blue stain for 30 min. Destaining was accomplished with deionized water overnight. After destaining, gels were imaged on a LAS-4000 (Fuji Film, Tokyo, Japan). Each gel lane was cut into 20 slices which were chopped further into ~1 mm³ pieces (Sup. Figure 1B, Supporting Information). Gel pieces were destained using three 20 min incubations in 50% ACN (v/v), 25 mM ammonium bicarbonate. Once all stain was removed, gel pieces were dehydrated in 100% ACN and then dried in a CentriVap (Labconco, Kansas City, MO) for 10 min or to complete dryness. Gel pieces were rehydrated in 50 mM ammonium bicarbonate containing 12.5 ng/µL trypsin on ice for 30 min, then incubated at 37 °C overnight. For extraction of peptides from the gel pieces, 100 µL 50% ACN, 5% formic acid was added and tubes were incubated while shaking at room temperature for 20 min. Tubes were centrifuged at 14 000× g for 5 min and supernatants transferred to clean tubes. These steps were repeated for a total of 3 extractions of peptides from the gel pieces. The extracts were taken to dryness in the CentriVap and the residues resuspended in mass spectrometry analysis buffer (5% ACN, 3% formic acid).

nano-LC–Mass Spectrometry Analysis and Protein Identification

Tryptic peptides from each of the gel slices were analyzed using a LTQ-Orbitrap hybrid mass spectrometer equipped with a nanoelectrospray source (Thermo Finnigan, San Jose, CA). Samples were loaded into a 180 µm × 20 mm nano-ACQUITY UPLC C₁₈ 5 µm trap column and then a 75 µm × 100 mm nano-ACQUITY UPLC 1.7 µm BEH analytical column and separated by a nano-ACQUITY UPLC system (Waters Corporation, Milford, MA). Peptides were separated at a flow rate of 500 nL/min. The solvent gradient of LC–MS was linear from 95% solvent A (0.1% formic acid in water) to 50% solvent B (0.1% formic acid in ACN) for 65 min. The eluent was introduced directly into an LTQ-Orbitrap mass spectrometer via nanoelectrospray ionization with PicoTip nanospray tip (New Objective, Woburn, MA). Each full MS scan was followed by eight MS/MS scans of the most intense ions with data-dependent selection using the dynamic exclusion option (Top 8 method). Thus, after the mass spectrometry, 20 “.dat” files were generated from each different fraction. Spectra were searched against the human IPI database v3.18 FASTA database (60 090 entries) using the SEQUEST search algorithm (SRF v. 5) of the Bioworks software v3.3.1sp1 (Thermo Fisher Scientific, San Jose) with the following parameters; parent mass tolerance of 10 ppm, fragment tolerance of 0.5 Da (monoisotopic), variable modification on methionine of 16 Da (oxidation), and maximum missed cleavage of 2 sites assuming the digestion enzyme trypsin. Search results were entered into Scaffold software (v2_05_01, Proteome Software, Portland, OR) for compilation, normalization, and comparison of spectral counts, etc. Protein identifications were made at the peptide probability of 95% and protein probability of 99% with at least two identified peptides. The shared and semitryptic peptides were excluded from spectral counts. Protein probabilities

and redundancy were assigned by the Protein Prophet algorithm.²⁷ Proteins that contained similar peptides and multiple isoforms, which could not be differentiated based on MS/MS spectra, were grouped into primarily assigned proteins. Moreover, for the estimation of false discovery rate (FDR), we searched against forward and reversed human protein databases using reverse mode by Bioworks. The protein list we are reporting here generated <3% FDR of peptide identification.²⁸ (FDR formula: $2 \times N_{\text{rev}} / [N_{\text{rev}} (\text{reverse database}) + N_{\text{real}} (\text{real database})]$)

Statistical Analysis of Mass Spectrometry Data Sets

The definition of spectral count is the total number of spectra for a protein identification and is an accepted label-free, semiquantitative value in proteomic analysis.²³ Since PLGEM provides signal-to-noise (STN) and p-values, which require the calculation of variance, PLGEM requires at least two replicates for the analysis. Spectral count data from duplicate analyses of nontumor and HCC samples were compared using a PLGEM in order to identify statistically significant protein changes between nontumor and HCC samples.^{24,25} PLGEM software was downloaded from www.bioconductor.org and imported into our web-based interface (<http://proteomics.carolinas.org/cgi-bin/plgem.pl>). While PLGEM was developed using a normalized spectral abundance factor (NSAF) as input, its performance with a limited number of replicates has been shown to improve when raw spectral count rather than NSAF is used.²⁹ Therefore raw spectral count was used as input in our PLGEM analysis. Estimated false discovery rates for PLGEM-generated significance lists were estimated using the Benjamini–Hochberg estimator.³⁰ The intersection between fractions were determined using a previously described linux-based intersect program.³¹

Western Blot

Twenty µg of subcellular fractionation product was loaded onto a 10% Bis-Tris NuPAGE gel and electrophoresed at 35 mA for 130 min at room temperature. Proteins were transferred to nitrocellulose membrane at 25 V at RT for 2 h using the Invitrogen Xcell II blot module (Invitrogen, Carlsbad, CA). Transfer efficiency (and in the case of fractionation products, evenness of sample loading) was confirmed by ponceau S staining of the membrane. Membranes were blocked for one hour at RT to overnight at 4 °C using 5% ECL blocking reagent in tris-buffered saline containing 0.1% Tween 20 (TBST). Membranes were incubated in primary antibody at 1:250–1000 dilution in blocking buffer overnight at 4 °C. Blots were washed 3 times with TBST for 5 min and then incubated in the appropriate secondary antibody conjugated to HRP for one hour at room temperature. Blots were then washed with TBST again 3 times, 5 min each. Chemiluminescent detection was accomplished using the ECL Advance kit and a Fuji LAS4000 digital imager.

Immunohistochemistry

Formalin-fixed, paraffin-embedded HCC tissue blocks for each patient were obtained from the pathology department at Carolinas Medical Center and four micrometer sections were prepared. Tissue array slides, catalog # LV809 (US Biomax, Rockville, MD) were deparaffinized in xylene and hydrated through graded alcohols to deionized water. The remainder of the procedure was performed using the AutostainerPlus. Endogenous peroxidase was blocked using 3% (v/v) H₂O₂ and slides were incubated for one hour with primary antibody at the following concentrations: Matrin 3, LETM1, and IQGAP2 at 1:50; ILF2 and fatty acid synthase at 1:100 dilution; OSF2 at 1:500. IgG from the corresponding species in which the antibody

Table 2. Top 5 Increased Subcellular Fraction Proteins with Significant Changes in HCC's vs Nontumor Bearing Livers

fraction ^a	IPI	identified proteins	GO localization ^b	NT ^c				HCC1 ^d	HCC2	STN ^e	p-value ^f
				NT1 ^c	NT2	HCC1 ^d	HCC2				
Cyto	IPI00014898	Isoform 1 of Plectin-1	unknown			70	79	434	413	19.28	0.0002
	IPI00105407	Aldo-keto reductase family 1 member B10	unknown		0	0	128	134	14.43	0.0002	
	IPI00294834	Aspartyl/asparaginyl beta-hydroxylase	cytoplasm, endoplasmic reticulum, intracellular organelle, membrane, organelle membrane, organelle part (integral to endoplasmic reticulum membrane)	0	0	107		96	12.13	0.0007	
	IPI00296337	Isoform 1 of DNA-dependent protein kinase catalytic subunit	unknown	12	10	107		126	10.58	0.0007	
	IPI00021812	Neuroblast differentiation-associated protein AHNAK (Fragment)	unknown	0	0	76		71	9.69	0.0011	
Mem	IPI00296337	Isoform 1 of DNA-dependent protein kinase catalytic subunit	unknown	4	5	100		130	10.85	0.0005	
	IPI00010180	Liver carboxylesterase 1 precursor	unknown	241	303	473		568	9.02	0.0007	
	IPI00105407	Aldo-keto reductase family 1 member B10	unknown	0	0	64		77	8.53	0.0007	
	IPI00220740	Isoform 2 of Nucleophosmin	intracellular organelle, organelle part (granular component), Nucleus (nuclear speck)	0	0	57		74	8.12	0.0011	
	IPI00022462	Transferrin receptor protein 1	cytoplasm (cytoplasmic membrane-bounded vesicle), endosome (endosome), extracellular region (extracellular region), intracellular organelle (endosome), membrane (integral to plasma membrane), plasma membrane (integral to plasma membrane)	4	3	66		66	7.59	0.0011	
Nuc	IPI00021304	Keratin, type II cytoskeletal 2 epidermal	unknown	1091	1765	1543		2474	9.36	0.0009	
	IPI00215965	heterogeneous nuclear ribonucleoprotein A1 isoform b	unknown	9	0	91		102	8.37	0.0010	
	IPI00386854	HNRPA2B1 protein	unknown	131	160	283		410	7.91	0.0010	
	IPI00104050	Thyroid hormone receptor-associated protein 3	intracellular organelle, nucleus, organelle part (mediator complex)	8	14	99		111	7.79	0.0010	
	IPI00296337	Isoform 1 of DNA-dependent protein kinase catalytic subunit	unknown	174	236	412		443	7.78	0.0010	
Cskel	IPI00007928	Pre-mRNA-processing-splicing factor 8	unknown	8	0	205		181	12.94	0.0002	
	IPI00420014	U5 small nuclear ribonucleoprotein 200 kDa helicase	unknown	0	0	134		122	10.90	0.0002	
	IPI00215638	ATP-dependent RNA helicase A	unknown	13	13	156		167	10.05	0.0002	
	IPI00217468	Histone H1.5	unknown	5	4	84		97	8.06	0.0004	
	IPI00221106	splicing factor 3B subunit 2	intracellular organelle, nucleus (nucleus)	0	0	76		75	7.91	0.0004	

^a Cyto (Cytosolic fraction), Mem (Membrane fraction), Nuc (Nuclear fraction), Cskel (Cytoskeletal fraction). ^b Gene Ontology localization by Scaffold program which uses the GO terms that appear in the NCBI database as part of a protein's description. ^c NT: Spectral counts (SC) of nontumor.

^d HCC: SC of HCC. ^e STN: Signal to noise. ^f P-value by PLGEM analysis.

was produced was used as a negative control. Secondary antibody was followed by peroxidase-conjugated streptavidin for 10 min

and 3'-diaminobenzidine for 5 min. Slides were removed from stainer, rinsed in water, counterstained with hematoxylin,

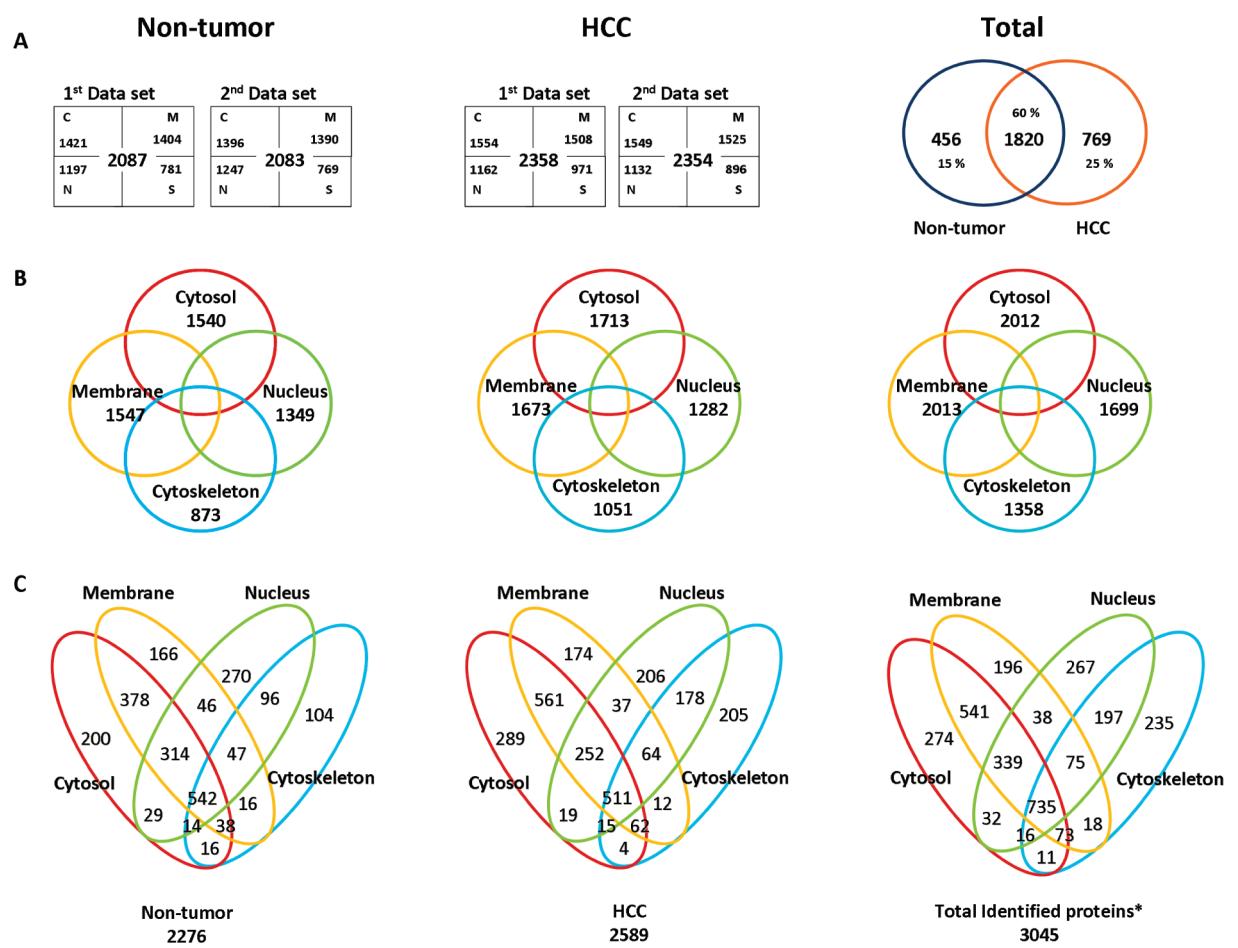


Figure 2. Distribution of proteins in nontumor liver and HCC of patient # 55. (A) Total intersection for duplicated data set for nontumor and HCC. (B) Intersection for each fraction. (C) Total intersection between nontumor and HCC.

dehydrated, cleared and mounted with resinous mounting media. Slides were imaged using an Olympus BX51 microscope equipped with an Olympus DP70 camera and DP controller imaging software (Olympus Corporation, Tokyo, Japan).

Immunofluorescence Staining of Human Liver Tissues and Cells

In immunofluorescence staining goat antimouse or rabbit antibody conjugated with Alexa 594 was used as a secondary antibody to detect the bound primary target (e.g., MATR3, ILF2, LETM1, and IQGAP2) and to amplify signals. Nuclear staining was performed by premixed 4',6-diamidino-2-phenylindole (DAPI) with proLong gold antifade mounting media whereas Pan-actin was used for cytoskeletal staining in the HCC tissue blocks and Huh-7 cell line. Fluorescence was visualized using a Zeiss LSM710 confocal microscope (Carl Zeiss Inc., Germany) with 10×, 20×, and 63× objectives. Image analysis was performed using Zeiss ZEN LE 2009 image analysis software. To allow direct comparisons, all images were captured using the same parameters.

RESULTS

Characteristics of Patients

The selected demographic, histologic, and biochemical features of subjects studied are listed in Table 1. Clinical

presentations were as follows:

Subject #23: A 60 year old man with a history of well controlled diabetes mellitus, otherwise in excellent health, presented with new onset gross hematuria. He was diagnosed with high-grade papillary bladder cancer and underwent CT scan of abdomen and pelvis as part of his staging workup. The abdominal CT scan revealed a 6 × 6.5 cm solid mass in the left hepatic lobe as well as a small lesion along the right posterior dome measuring 2.2 cm in diameter. Serum alpha fetoprotein level was found to be approximately 26 000 ng/mL, and a core needle biopsy subsequently confirmed the presence of hepatocellular carcinoma (HCC). He did not have any history of chronic liver disease or viral hepatitis and denied any history of encephalopathy, jaundice, or ascites.

Subject #55: A 76 year old man who had possible NAFLD in the setting of diabetes and hyperlipidemia, heterozygous for the C282Y gene mutation of HFE, presented to his primary care physician complaining of right shoulder pain diagnosed to have been due to bursitis. He was prescribed an over the counter NSAID, and a few weeks later presented with gross hematuria. CT scan of the abdomen and pelvis revealed a <5 mm cyst in the superior aspect of the right kidney which was thought to have been benign and a liver mass 0.5 × 12.5 cm in greatest dimension. His hematuria resolved upon discontinuing the aforementioned anti-inflammatory medication.

Table 3. Selected Biosignature Candidate Proteins Identified in the Nuclear Fraction with Significant Changes in HCC's vs Nontumor Bearing Livers

no.	ID	nucleus protein	GO localization ^a	patient 23		patient 55		patient 65	
				STN ^b	p-value ^c	STN	p-value	STN	p-value
1	IPI00002459	annexin VI isoform 2	Cytoplasm (perinuclear region of cytoplasm), membrane and plasma membrane (apical plasma membrane).	-0.9	0.0324	-5.4	0.0010	-15.1	0.0000
2 ^d	IPI00004101	Betaine--homocysteine S-methyltransferase	Unknown	-2.7	0.0016	-4.2	0.0018	-6.4	0.0000
3 ^d	IPI00005198	Interleukin enhancer-binding factor 2	intracellular organelle. nucleus, organelle part (nucleolus)	0.7	0.0568	5.3	0.0023	4.9	0.0000
4	IPI00006196	Isoform 2 of Nuclear mitotic apparatus protein 1	cytoskeleton and intracellular organelle (spindle pole), nucleus and organelle part (nuclear matrix)	0.5	0.1244	5.5	0.0023	2.3	0.0016
5 ^d	IPI00006663	Aldehyde dehydrogenase. mitochondrial precursor	cytoplasm and intracellular organelle and mitochondrion (mitochondrion)	-1.4	0.0122	-6.1	0.0008	-9.3	0.0000
6	IPI00007940	SPFH domain family, member 1	Unknown	-2.9	0.0011	-2.4	0.0049	-2.5	0.0016
7	IPI00012912	Carnitine O-palmitoyltransferase 2. mitochondrial precursor	cytoplasm and intracellular organelle and mitochondrion (mitochondrion)	-4.9	0.0002	-1.8	0.0096	-4.4	0.0000
8	IPI00017297	Matrin-3	Nuclear inner membrane	0.2	0.2210	1.9	0.0149	6.8	0.0000
9 ^d	IPI00017592	Leucine zipper-EF-hand-containing transmembrane protein mitochondrial precursor	Cytoplasm (mitochondrion). Membrane and Plasma membrane (integral to plasma membrane)	1.2	0.0185	1.2	0.0388	7.0	0.0000
10	IPI00025276	Isoform XB of Tenascin-X precursor	extracellular region (proteinaceous extracellular matrix)	-2.8	0.0014	-9.9	0.0004	-1.8	0.0049
11 ^d	IPI00025307	Cytochrome P450 27. mitochondrial precursor	Unknown	-3.1	0.0010	-7.7	0.0004	-4.9	0.0000
12	IPI00026089	Splicing factor 3B subunit 1	cytoskeleton. intracellular organelle. nucleus, organelle part (spliceosome)	0.5	0.1244	2.1	0.0134	4.8	0.0000
13 ^d	IPI00026781	Fatty acid synthase	cytoplasm (glycogen granule)	-9.5	0.0000	-4.7	0.0015	-8.8	0.0000
14 ^d	IPI00027230	Endoplasmin precursor	Unknown	-3.0	0.0010	-5.2	0.0011	-4.1	0.0002
15 ^d	IPI00027834	heterogeneous nuclear ribonucleoprotein L isoform a	intracellular organelle. nucleus organelle part (nucleoplasm)	1.0	0.0353	3.6	0.0041	5.8	0.0000
16 ^d	IPI00028031	Isoform 1 of Very-long-chain specific acyl-CoA dehydrogenase. mitochondrial precursor	Unknown	-1.7	0.0060	-4.4	0.0016	-4.2	0.0000
17	IPI00030363	Acetyl-CoA acetyltransferase. mitochondrial precursor	Unknown	-2.4	0.0023	-1.4	0.0157	-5.0	0.0000
18	IPI00031131	Adipocyte plasma membrane- associated protein	Unknown	-3.5	0.0007	-1.8	0.0090	-3.1	0.0007
19	IPI00031545	Isoform Long of Inositol 1,4,5-trisphosphate receptor type 2	Unknown	-0.9	0.0322	-6.0	0.0008	-1.0	0.0271
20 ^d	IPI00217872	Phosphoglucomutase 1	cytoplasm	-1.5	0.0102	-2.7	0.0038	-2.1	0.0028
21 ^d	IPI00299048	IQ motif containing GTPase activating protein 2	cytoskeleton. intracellular organelle (actin cytoskeleton)	-3.9	0.0006	-7.5	0.0005	-7.0	0.0000

^a Gene Ontology localization by Scaffold program which uses the GO terms that appear in the NCBI database as part of a protein's description. ^b STN: Signal to noise. ^c P-value by PLGEM analysis. ^d Proteins published previously in association with HCC.

Subject #65: A 66 year old lady with history of supraventricular tachycardia (SVT), COPD, breast CA, status-post lumpectomy and chemo-radiotherapy, was in her usual state of health until she presented to the emergency room with precordial and

epigastric pain. Her symptoms were associated with intermittent nausea, but she denied any dyspnea on exertion, shortness of breath, vomiting or diaphoresis. Her pain did not respond to nitrates but was relieved by morphine. She had an extensive

cardiac workup with unremarkable results, including a nuclear stress test. She was not thought to be in COPD exacerbation, nor did she have evidence of pulmonary embolism by CT angiography. As part of her evaluation, an abdominal MRI revealed a lobular hepatic contour with enhancing masses involving segments 2, 3, 4A and 4B. The confluent mass involving segments 2 and 3 measured 9.8×6.9 cm and was worrisome for HCC. Finally, there was a thrombus occluding both the left and right portal vein along with its anterior segmental branch. Exploratory surgery confirmed HCC which was resected. The hepatic parenchyma around the tumor was found to be unremarkable (Table 1).

Protein Identification for Subcellular Fractionation

We performed subcellular fractionation and proteomic analysis of the No. 55 patient (55-NT/-HCC) to identify differentially expressed and translocated proteins from NT and HCC tissue. We analyzed two technical replicates from four fractions using an Orbitrap-LTQ mass spectrometer for statistical evaluation by PLGEM as described in Figure 1 (see also Experimental Procedures). A total of 2 174 504 MS/MS spectra were generated from all 8 fractions (cytosolic, membrane, nuclear and cytoskeletal fractions of NT and HCC tissues (Sup. Table S1a–d, Supporting Information); 3045 unique proteins from 8 fractions with <3% false discovery rate (FDR) were identified (Sup. Table S2, Sup. Figure 8, Supporting Information). FDRs for PLGEM-generated significance lists were estimated using the Benjamini-Hochberg estimator.³⁰ Though the PLGEM does not provide an FDR, the user can estimate the FDR of protein by inputting a significance level of *p*-value. The number of false positives from simultaneous protein comparison is below 0.1%, giving an estimated false discovery rate maximum below 4.5% for the putative list of differentially expressed nuclear proteins. In detail, the expected number of nuclear fraction false positives was 0.001 (*p*-value) \times 1699 (total proteins) = 1.69, giving an estimated FDR of 1.69/38 (the number of protein with *p*-value <0.001) = 4.5% for the putative list. The estimated FDR of PLGEM analysis in different levels of *p*-values was shown in Sup. Figure 8 (Supporting Information). Finally, we found 141 putative unique proteins for the HCC tissue of patient No. 55. Of these, 48 (FDR: 4.2%) were identified in cytosol, 55 (FDR: 3.6%) in membrane, 38 (4.5%) in nucleus and 67 (2.1%) in cytoskeleton. Table 2 lists the most increased proteins found in HCC tissues.

We next classified common and unique proteins from a total of 3045 proteins that were distributed in all NT and HCC fractions using our previously described “intersect” program. 2276 were identified in nontumor proteins, 2589 proteins were identified in tumor among duplicate data (Figure 2C). Among them, 456 proteins (15%) were identified unique to NT, 769 (25%) were unique to HCC, and 1820 proteins (60%) were common to both (Figure 2A). Protein fractions of each duplicate data set and the numbers of unique proteins identified specifically in NT and HCC fractions are illustrated as described in Figure 2B and C. The gene ontology (GO) analysis was performed using Scaffold GO annotations function as described (Sup. Figure 5, Supporting Information); the GO localization for selected proteins is presented in Tables 2 and 3.

Fractionation Efficiency

In this experiment, reproducibility was demonstrated by the overlap of 80% in technical replicate data from patient No. 55 subcellular fractions. We also compared the fractionation efficiency of the four fractionation buffers. We found that the overall trend in protein identification was quite similar among the four fractions. The overlap of protein identifications in cytosol and

membrane was 68%, membrane and nucleus showed overlap of 44%, and nucleus and cytoskeleton overlapped by 46% (Sup. Figure 2 and 7, Supporting Information). This result indicates that nuclear and cytoskeletal fraction buffers preferentially demonstrated a lower degree of carry over between fractions and therefore represent the most efficient fractionation for determining localization.

Targeted Subcellular Fractionation

Based on the analysis of all four fractions from patient 55 NT and HCC tissues, we concluded that the nuclear and cytoskeletal fractions demonstrated the highest fractionation efficiency. In order to identify more reliable biosignatures for HCC, we performed not only technical replicate proteomics profiling analysis but performed biological replicate experiments with tissues from patients 23 and 65 and specifically targeted the nuclear and cytoskeletal fractions. The NT and HCC tissues of patients 23 and 65 were fractionated using the above method. The targeted nuclear and cytoskeletal fractions were then analyzed by LC–MS/MS proteomics. We identified 2575 nonredundant proteins (Sup. Figure 3A, Supporting Information) from the nuclear fractions of all three patients combined, among which 827 were common to all three patients, 392 proteins were found only in nontumor liver, and 638 proteins were found only in the HCC tissues (Sup. Figure 3C, Supporting Information). In the cytoskeletal fraction, 2043 proteins were identified from three biological samples and 590 proteins (Sup. Figure 3B, Supporting Information) were common to all three patients, whereas 327 and 672 proteins (Sup. Figure 3D, Supporting Information) were unique to nontumor liver and HCC tissues, respectively (Sup. Table S1c–h, Supporting Information).

Data-mining for Molecular Signature of HCC

To examine the statistical relevance of regulated proteins from the targeted subcellular fractions, we transferred each data set to PLGEM analysis and generated signal-to-noise (STN) levels and *p*-values (Sup. Table S3 and S4, Supporting Information). In the next step, we selected primary candidate proteins for molecular signature which were commonly up-regulated or down-regulated; 305 nuclear and 211 cytoskeletal proteins were categorized as the primary candidates in these fractions. These proteins were intersected with the HUPO plasma database (3020 proteins with two or more peptides)^{32,33} in an effort to identify potentially secreted proteins that were differentially expressed in cancer (Table 3). As a result, final 21 candidates were determined by PLGEM *P*-value of <0.05 (patient No. 55). We chose for validation target proteins that were common in all three patients and also present in the HUPO plasma database. Although ILF2 was not included in the HUPO plasma database, ILF2 was selected for further validation due to its increase among all three patients as well as previously reported up-regulation at the gene level in HCC.^{34,35} In addition to the tissue samples, we further applied stringent proteomic analysis of the human hepatoma cell lines, Huh7.5 and CON1 (data not shown here) to extend the confidence of protein identification of liver tissue in an established cell line system. The spectral counts of the selected candidates are plotted on Figure 3. The proteins IQGAP2 and FAS were decreased, while LETM1, MATR3 and NUMA were increased proteins. The spectral counts of all fractions for IQGAP2 were decreased in the HCC fractions along with the lower counts of Huh7.5 and CON1. Other proteins in Figure 3 showed a similar pattern in the Huh7.5 and CON1 cell lines.

Validation using Western Blot

We next validated the identification and semiquantification results of proteomic analysis using antibodies against eight of the

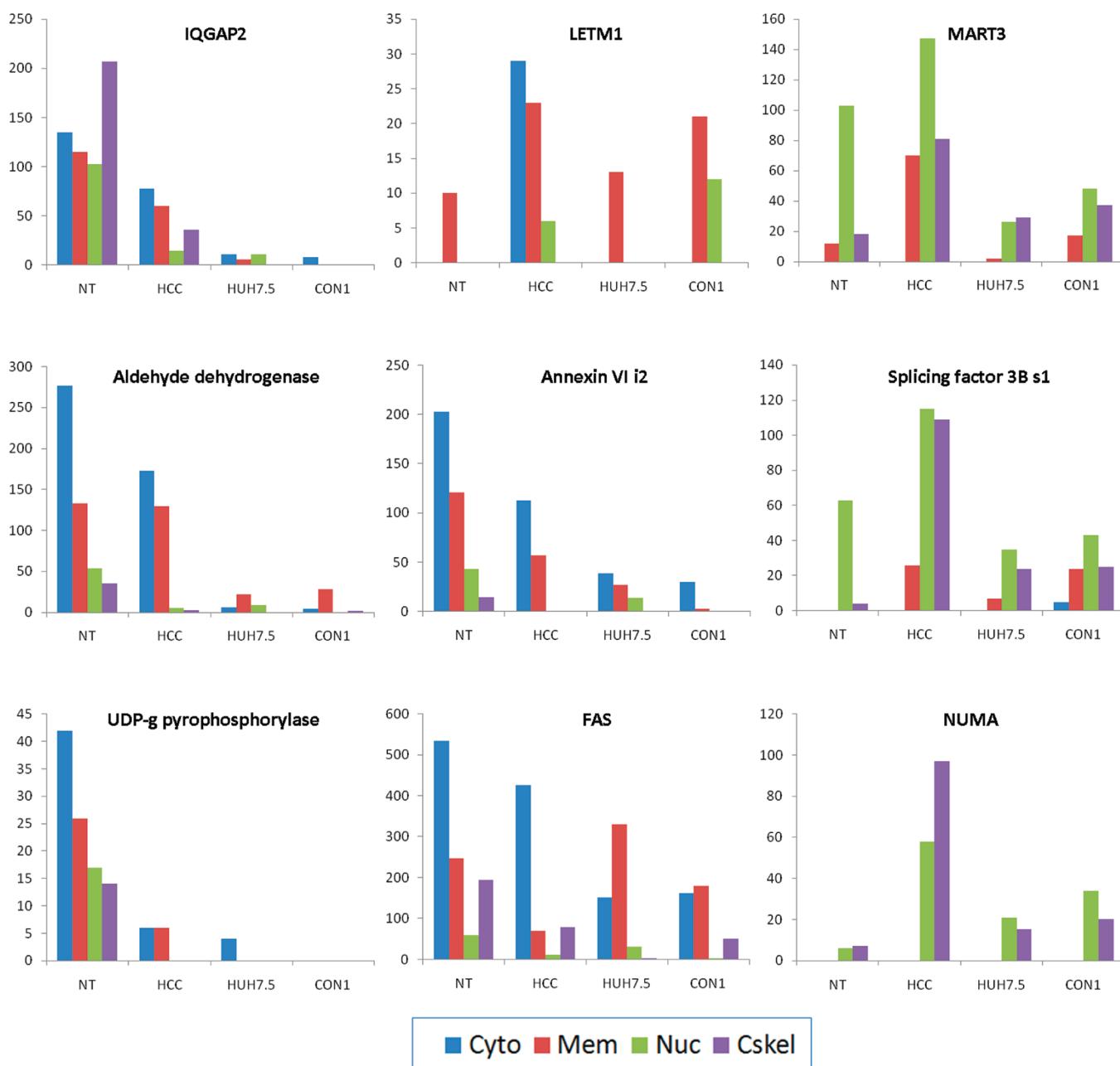


Figure 3. Selected common HUPO proteins between human tissue and HCC cell lines. The subcellular fractionations were performed with nontumor (NT), HCC tissue of Patient No. 55, Huh7.5, CON1 cell lines. The mean spectral counts (duplicate data set) for nontumor (NT) and HCC and spectral counts (single data set) were plotted.

biosignature candidates with subcellular fractions from nontumor liver and HCC tissues. Our goal was to compare the identification and quantification results from the LC–MS/MS analysis with the results from the Western blotting, comparing eight proteins in all four fractions. Through these validation studies we identified HCC associated up-regulation in proteins from nucleus and cytoskeleton fractions specifically, LETM-1, OSF2, NUMA and MATR3, and down-regulation of IQGAP2, FAS, CPS1 and TRAP150. Results indicated that MATR3 increased 2.1 fold, LETM1 and ILF increased 1.2 fold in nuclear fraction, while FAS and IQGAP in cytoskeletal fraction decreased 2.4 and 7.2 fold, respectively (Figure 4, left). In most cases, Western blotting data were in good agreement with spectral count data, thus

confirming higher expression for LETM-1, OSF2, NUMA and MATR3 and lower expression for IQGAP2, FAS, CPS1 and TRAP (Figure 4, left and Sup. Figure 4, Supporting Information).

Validation using Immunohistochemistry and Immunofluorescence Analysis

Immunohistochemical validation was also performed to examine the expression level and localization in nontumor and HCC tissues. Figure 4 shows representative immuno-histochemistry (IHC) staining of patient tissue as well as Western blot results for target proteins in selected patient samples. Immunostaining verified the expression of these proteins within the tissues. MATR3 (A), LETM1 (B) and ILF2 (C) (increased proteins) were identified in

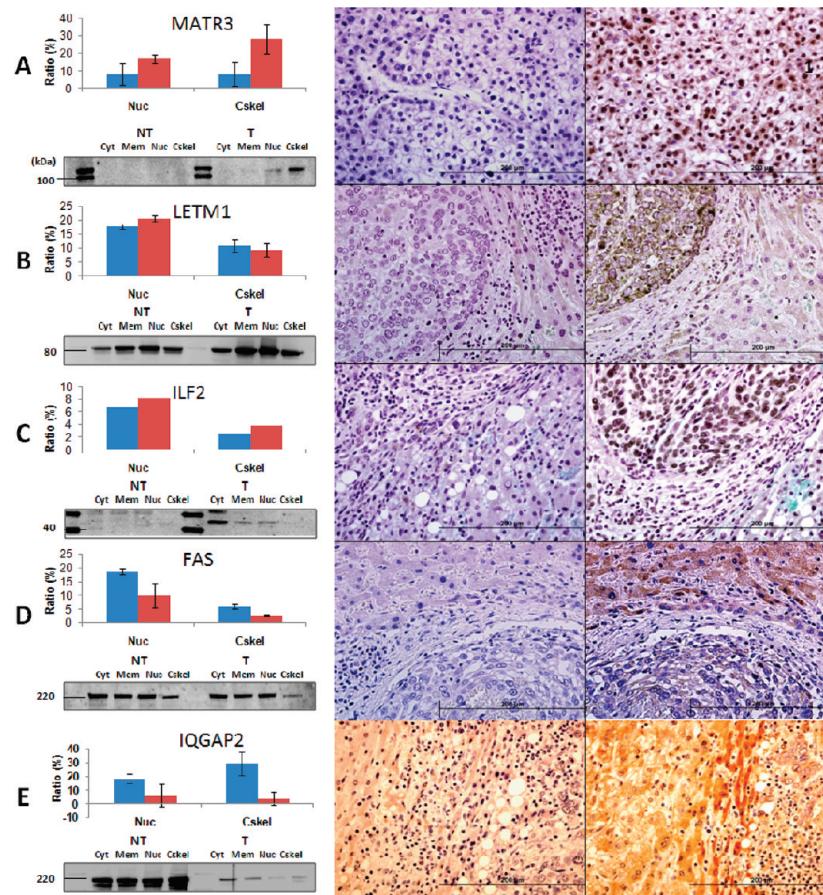


Figure 4. Western blot and Immunohistochemical validation of biosignature candidates in HCC patients. A, MATR3; B, LEMT1; C, ILF2; D, FAS; E, IQGAP2. Left, Western blot quantitative analysis of representative patients subcellular fractionation (Upper, Quantitative analysis $n = 2$, SD; Lower, Western image). Middle and Right, Representative photographs of immunohistochemical staining of formalin-fixed, paraffin-embedded liver tumor tissue with same antibody as western (Right) and no antibody control (Middle). Magnification: Middle and Right panel, $400\times$. Cyt, Cytosolic fraction; Mem, membrane fraction; Nuc, Nuclear fraction; Cskel, Cytoskeletal fraction.

nucleus while FAS (D) and IQGAP (E) (decreased proteins) were identified in the cytoskeletal region. Careful analysis of IHC revealed that ILF2 was expressed in the nucleus and outer nucleus regions, while the perinuclear region was stained with LEMT1 antibody. Results showed that the HCC tissues expressed more of these proteins compared with NT. For FAS, decreased protein expression was clearly found in the NT areas like cytoskeleton. IQGAP2 was detected more clearly in the fibrous septa. Especially LEMT1, reported to be located in cytoplasm and membrane by GO analysis (Table 3, Sup. Figure 5, Supporting Information), appeared to be translocated to perinuclear region according to our immunohistochemistry analysis of three patients. These results show localization by subcellular fractionation and proteomic analysis using spectral count are well matched according to validation by IHC.

To investigate the expression level of the putative biosignatures in different HCC grades and compare them between normal and HCC patients, we utilized a commercial tissue array including normal tissue as well as various grades and stages of HCC (Figure 5). In the spectral counts, signal-to-noise (STN) of MATR3 was increased alongside tumor grade (Table S3, Supporting Information). IHC analysis (Figure 5A) showed MATR3 expression in the nucleus (Figure 5A). ILF2 was found in the nuclear region as well (Figure 5B). According to western analysis and IHC of patient tissue blocks, LEMT1 was mainly increased in membrane and nuclear

fractions (Figure 4B). The tissue array results showed LEMT1 existed in peripheral nuclear and extra-nuclear region (Figure 5Ca–f). The histograms in Figure 5Ag, Bg and Cg are representative of the staining of the nuclear region magnified below the table. The intensities of the staining for the associated antibodies were analyzed using the ImageJ program. LEMT1 showed higher expression levels near the outer and inner regions of the nucleus (Figure 5Cg) compared to MATR3 (Figure 5Ag). For further verification of translocation, we performed immunofluorescent analysis for MATR3 (Figure 5Ah–Aj) ILF2 (Figure 5Bh–Bj) and LEMT1 (Figure 5Ch–Cj). To confirm and visualize localization of candidate proteins, immunofluorescent staining was performed in the Huh-7 cell line. MATR3 (Figure 5Ai) and ILF2 (Figure 5Bi) were imposed with nuclear marker DAPI (Figure 5Ah and Bh), especially LEMT1 was highly expressed in perinuclear space (Figure 5Ch–j). The combined images of target proteins stained with cytoskeletal marker Pan-actin additionally showed characteristics of localization. These immunofluorescent findings were well matched with Western blotting and IHC results (Sup. Figure 6, Supporting Information).

DISCUSSION

Normalization to correct for differences is common in proteomics analysis. Fluctuations in total spectral count are often observed between technical replicates of the same sample.

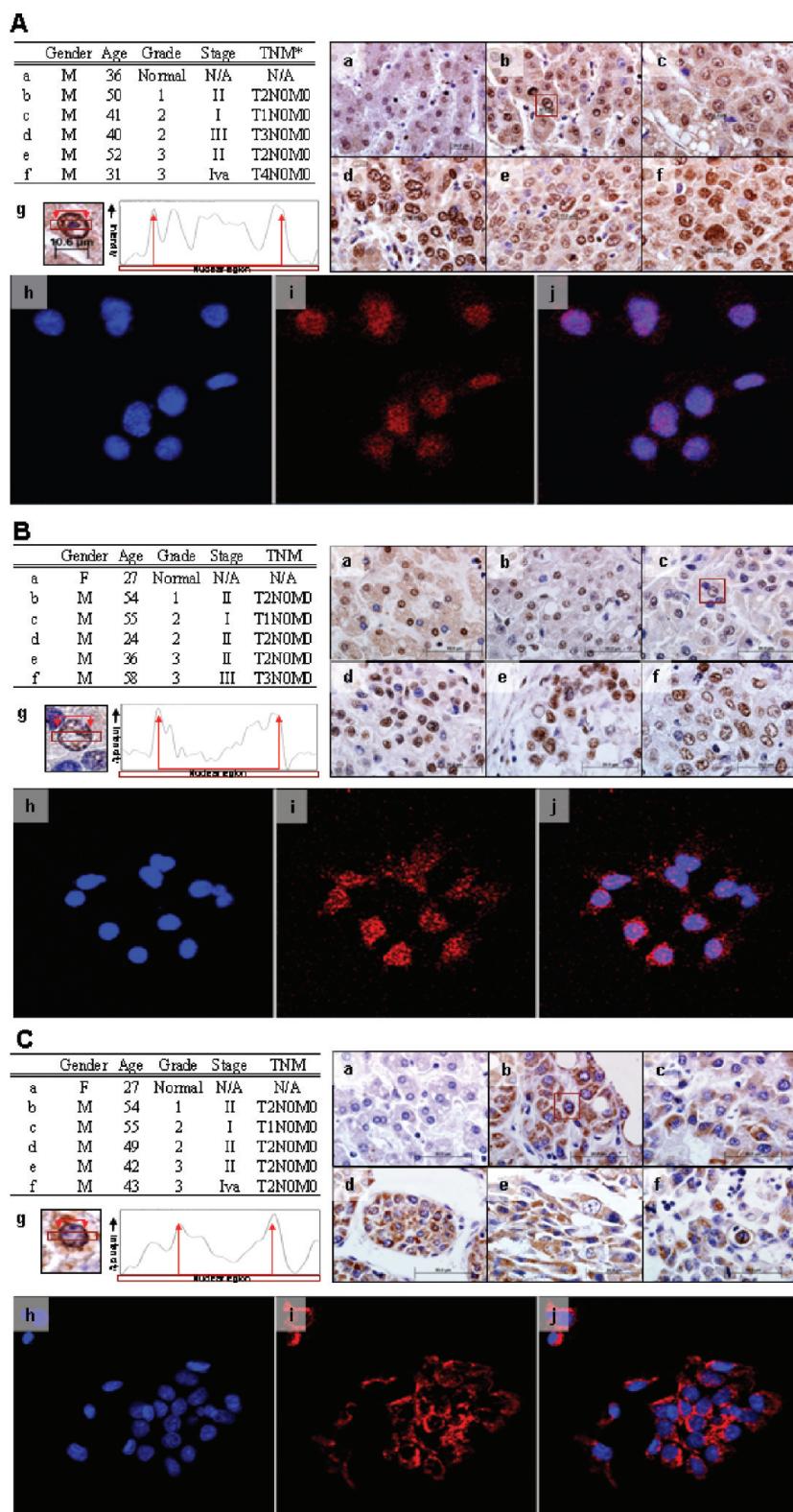


Figure 5. Immunohistochemical and immunofluorescence validation of biosignature candidates for HCC tissue array and Huh-7 cell line. (A) MATR3. (B) ILF2. (C) LETM1. Tables show tissue block information of HCC patients (a–f). (g left) Representative nuclear region (for A, B, and C). (g right) Nuclear intensity. Immunofluorescence validation was performed with Huh-7 cell line. (h) DAPI (nuclear marker); (i) target protein; (j) : integrated image (h)+(i). Magnification: 200×. *TNM: TNM grading (T, primary tumor; N, regional lymph nodes; M, distant metastasis).

Nonetheless, it is necessary to consider whether normalization is essential for a reliable result. One normalization method is

known as the normalized spectral abundance factor (NSAF), where NSAF is spectral count divided by sequence length and

normalized by total spectral count.²⁵ Recently, we investigated the use of NSAF versus raw spectral count in the application of PLGEM to a two-replicate data set. We determined that the performance of PLGEM using raw spectral counts equals or exceeds that of NSAF.²⁹ In addition, we compared the effect of normalization on our experimental results in this study, by employing the normalization function, “quantitative view”, within the Scaffold program. We applied this normalization to the data generated in the nuclear fraction of patient # 55. The software calculates normalized spectral count by averaging the spectral counts across all categories and then multiplying the spectrum counts. This data set was transferred to PLGEM analysis. During the course of this study, we assigned statistical significance to PLGEM generated STN values with $p < 0.05$. At this level, less than 1% of proteins demonstrated a change in differential expression when using the Scaffold normalized spectral count data as opposed to raw spectral count. As this represents only a small fraction of the reported data, and in light of our previous studies indicating better performance using raw spectral count, we concluded the PLGEM values using raw spectral are comparable to that of normalized spectral count.

Many digestion protocols have recommend alkylation of the protein to avoid reformation of disulfide bonds. A study showed that there was no noticeable difference in heterogeneity of peptides that contained cysteine, although sometimes alkylated samples showed slightly higher yields than unalkylated samples.³⁶ One popular in-gel digestion protocol was introduced in 1996.³⁷ This protocol was further optimized and reported in 2006.³⁸ The critical step in this protocol is to decrease the risk of contaminating samples with human keratins because this might mask real data analysis in interesting samples. As the number of steps are added to a protocol, the probability of keratin contamination will be increased. Even though an additional reduction and alkylation step is performed in silver-staining or *de novo* sequencing samples,³⁹ skipping the reduction and alkylation steps is recommended for rapid identification using Coomassie staining.³⁸ Due to these factors, our protocol (following that of Mann's group) eliminates unnecessary, potentially contaminating steps. As a result, we searched our data using only methionine oxidation.

This study, which included subcellular fractionation and LC–MS/MS analysis of HCC tissues as well as statistical modeling in HCC versus nontumor bearing liver, resulted in the identification of 3045 unique proteins in HCC from 1 patient (protein probability >99%, at least 2 peptides). In addition, we performed cellular compartment targeted biological replicate experiments on samples from 3 patients. Combined, nonredundant data from the three patients identified 2575 proteins in the nuclear fraction and 2043 proteins in the cytoskeletal fraction (Sup. Figure 3A–D, Supporting Information).

A total of 399 proteins were up- or down-regulated in the two fractions of all three patients. Of these, 188 were unique to the nucleus, and 94 were unique cytoskeletal proteins. The intersection of these 399 proteins and the HUPO database selected 21 proteins which have been identified in nucleus and cytoskeleton. The increased expression and potential translocalization of some of these proteins may indicate a novel modulation of HCC. Overall, we identified 21 potential biosignature candidates in the nuclear and cytoskeletal fractions using PLGEM values (cut off: $p < 0.05$ of patient # 55). Of these, we validated MATR3, LETM1, IQGAP2, Fatty acid synthase, and OSF2, both by Western blotting and immunohistochemistry. Furthermore, MATR3, ILF2 and LETM1 were evaluated with immunohistochemistry

analysis using a multitissue array and immunofluorescent microscopy localization experiments. Interestingly, LETM1 appears to be translocated to the peripheral nuclear region in HCC tissues.

It has been postulated that MATR3, an abundant protein of the internal nuclear matrix, may play a role in transcription or may interact with other nuclear matrix proteins to form the internal fibrogranular network.⁴⁰ Defects of MATR3 are known to be the cause of myopathy distal type 2.⁴¹ Recently, Roy et al. compared the protein composition of transitional endoplasmic reticulum (tER) isolated from liver tumor nodules of aflatoxin-treated rat models with that of control liver. Proteins known to be involved in transcription including MATR3, increased in tumor tER relative to that found in control tER.⁴² In the current study, the spectral counts of MATR3 were increased in all three patients, based on PLGEM analysis. When evaluating gene ontology analysis, MATR3 was confirmed as a nuclear protein. In the next step, we validated this protein with Western blotting and IHC. We confirmed that the protein was up-regulated not only in western analysis but also by immunohistochemistry in the nuclear fraction ($n = 3$). Also of particular interest in IHC experiments was that, as tumor grade advanced, the size of nuclei (with staining of MATR3) was increased using an HCC tissue array sample with 80 patients (Figure 5). Furthermore, we verified MATR3 localization in the nuclear region via immunofluorescent staining. Therefore, it is suggested that MATR3 represents a possible biosignature molecule according to these results (Figure 5).

The LETM1 gene was originally identified as one of the genes that are chromosomally deleted in patients with Wolf-Hirschhorn syndrome.⁴³ This gene is important for the maintenance of mitochondrial tubular networks and for the assembly of the supercomplexes of the respiratory chain.⁴⁴ It was reported that subcellular localization of LETM1 is mitochondrion inner membrane.^{44–46} Interestingly, our study demonstrated that LETM1 is up-regulated in the tumor nuclear fraction by western analysis. IHC showed this protein to be translocated to the peripheral nuclear region as well as the outer nucleus. We confirmed the translocation of LETM1 not only by IHC of an 80 patient HCC tissue array but also by immunofluorescent staining experiments with our three patient tissue blocks.

It has been reported that HepG2, Huh-7 (pediatric hepatoblastoma cell lines), and Hep3B (adult HCC cell line), express high levels of IQGAP2 mRNA, whereas SK-HEP-1 (adenocarcinoma cell line) shows no detectable expression.⁴⁷ Sun et al. reported that AFP-positive HepG2 and Hep3B expressed high levels of IQGAP2 protein. However these levels were not detected in AFP-deficient liver cancer cells as well as in normal liver cells (HL-7702). They suggested that IQGAP2 might play a significant role in hepatocarcinogenesis, especially for AFP positive HCC.¹⁴ Schmidt et al. provided evidence that loss of IQGAP2 expression leads to development of HCC in mice.⁴⁸ In our study, IQGAP2 protein was decreased in all patients. These results support the previously reported data in our human HCC tissues.

In combination with these previously reported observations, our results suggest that MATR3 and LETM1 represent molecular biosignature candidates with potential as therapeutic targets based on their association with the pathological processes of HCC.

Future studies will further validate preliminary data showing differential expression of biosignature proteins quantified in the plasma of HCC patients compared to healthy control subjects. Several proteins listed in our biosignature candidate panel have already been identified in plasma and our next study will be

designed to determine quantitative and analytical validation and to facilitate and optimize protocols for blood proteomics in HCC. We also plan to characterize the proteome of HCC in patients with known underlying chronic hepatitis B, hepatitis C, and cirrhosis due to alcoholic or nonalcoholic fatty liver disease.

CONCLUSIONS

Subcellular tissue proteomics identified 3045 proteins of nontumor and HCC tissue from cytosolic, membrane, nuclear, and cytoskeletal fractions. The final lists of highly differentiated proteins were searched for potentially translocated proteins in HCC from soluble compartments to the nuclear or cytoskeletal compartments. The current work refined 21 potential targets of HCC and validated the potential molecular targets, MATR3, LETM1, ILF2 and IQGAP2 by Western blotting, immunohistochemistry, and immunofluorescent staining.

ASSOCIATED CONTENT

Supporting Information

Supplementary figures and tables. This material is available free of charge via the Internet at <http://pubs.acs.org>.

AUTHOR INFORMATION

Corresponding Author

*Sun-Il Hwang, Ph.D., Carolinas Medical Center, Proteomics Laboratory for Clinical and Translational Research, 1000 Blythe Blvd, Charlotte, NC, 28203. Phone: 704-335-9642. E-mail: sunil.hwang@carolinas.org.

ACKNOWLEDGMENT

We thank Ms. Roni Staten, Ms. Gale Groseclose, and the members of the Liver-Biliary-Pancreatic Repository, Dr. Qi Long Lu in the McColl-Lockwood Muscular Dystrophy Laboratory and Ms. Jane Ingram and Dr. Helen Gruber in the Histology Core Laboratory at Carolinas HealthCare System. This research was supported by a CHS Cannon Research Grant (CRG 08-026).

ABBREVIATIONS

AFP, alpha-feto-protein; CPS1, carbamoyl-phosphate synthase [ammonia], mitochondrial; Cskel, cytoskeleton fraction; CEB, cytosol extraction buffer; Cyt, cytosol fraction; FAS, fatty acid synthase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; HCC, hepatocellular carcinoma; HUPO, Human Proteome Organization; IHC, Immunohistochemistry; ILF2, interleukin enhancer-binding factor 2; LETM1, leucine zipper-EF-hand-containing transmembrane protein 1 and EF-hand domain-containing protein 1, mitochondrial; MATR3, matrin3; MEB, membrane extraction buffer; Mem, membrane/particulate fraction; NT, nontumor; NEB, nuclear extraction buffer; Nuc, nuclear fraction; NUMA, nuclear mitotic apparatus protein 1; OSF2, periostin; PLGEM, Power Law Global Error Model; IQGAP2, Ras GTPase-activating-like protein IQGAP2; STN, signal to noise; SILAC, stable isotope labeling with amino acids in cell culture; TRAP150, thyroid hormone receptor-associated protein 3; tER, transitional endoplasmic reticulum; tbst, Tris-buffered saline containing Tween 20; UPLC, ultra performance liquid chromatography

REFERENCES

- (1) Ferenci, P.; Fried, M.; Labrecque, D.; Bruix, J.; Sherman, M.; Omata, M.; Heathcote, J.; Piratsivuth, T.; Kew, M.; Otegbayo, J. A.; Zheng, S. S.; Sarin, S.; Hamid, S. S.; Modawi, S. B.; Fleig, W.; Fedail, S.; Thomson, A.; Khan, A.; Malfertheiner, P.; Lau, G.; Carillo, F. J.; Krabshuis, J.; Le Mair, A. Hepatocellular carcinoma (HCC): a global perspective. *J. Clin. Gastroenterol.* **2010**, *44* (4), 239–45.
- (2) El-Serag, H. B.; Rudolph, K. L. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. *Gastroenterology* **2007**, *132* (7), 2557–76.
- (3) Gish, R. G. Hepatocellular carcinoma: overcoming challenges in disease management. *Clin. Gastroenterol. Hepatol.* **2006**, *4* (3), 252–61.
- (4) Trauner, M.; Arrese, M.; Wagner, M. Fatty liver and lipotoxicity. *Biochim. Biophys. Acta* **2010**, *1801* (3), 299–310.
- (5) Jemal, A.; Siegel, R.; Ward, E.; Hao, Y.; Xu, J.; Murray, T.; Thun, M. J. Cancer statistics, 2008. *CA Cancer J. Clin.* **2008**, *58* (2), 71–96.
- (6) Poon, R. T.; Ng, I. O.; Fan, S. T.; Lai, E. C.; Lo, C. M.; Liu, C. L.; Wong, J. Clinicopathologic features of long-term survivors and disease-free survivors after resection of hepatocellular carcinoma: a study of a prospective cohort. *J. Clin. Oncol.* **2001**, *19* (12), 3037–44.
- (7) Chignard, N.; Beretta, L. Proteomics for hepatocellular carcinoma marker discovery. *Gastroenterology* **2004**, *127* (5 Suppl 1), S120–5.
- (8) Yuen, M. F.; Cheng, C. C.; Lauder, I. J.; Lam, S. K.; Ooi, C. G.; Lai, C. L. Early detection of hepatocellular carcinoma increases the chance of treatment: Hong Kong experience. *Hepatology* **2000**, *31* (2), 330–5.
- (9) Beretta, L. Comparative analysis of the liver and plasma proteomes as a novel and powerful strategy for hepatocellular carcinoma biomarker discovery. *Cancer Lett.* **2009**, *286* (1), 134–9.
- (10) Daniele, B.; Bencivenga, A.; Megna, A. S.; Tinessa, V. Alpha-fetoprotein and ultrasonography screening for hepatocellular carcinoma. *Gastroenterology* **2004**, *127* (5 Suppl 1), S108–12.
- (11) Lee, N. P.; Chen, L.; Lin, M. C.; Tsang, F. H.; Yeung, C.; Poon, R. T.; Peng, J.; Leng, X.; Beretta, L.; Sun, S.; Day, P. J.; Luk, J. M. Proteomic expression signature distinguishes cancerous and nonmalignant tissues in hepatocellular carcinoma. *J. Proteome Res.* **2009**, *8* (3), 1293–303.
- (12) Codarin, E.; Renzone, G.; Poz, A.; Avellini, C.; Baccarani, U.; Lupo, F.; di Maso, V.; Croce, S. L.; Tiribelli, C.; Arena, S.; Quadrifoglio, F.; Scaloni, A.; Tell, G. Differential proteomic analysis of subfractionated human hepatocellular carcinoma tissues. *J. Proteome Res.* **2009**, *8* (5), 2273–84.
- (13) Chaerkady, R.; Harsha, H. C.; Nalli, A.; Gucek, M.; Vivekanandan, P.; Akhtar, J.; Cole, R. N.; Simmers, J.; Schulick, R. D.; Singh, S.; Torbenson, M.; Pandey, A.; Thuluvath, P. J. A quantitative proteomic approach for identification of potential biomarkers in hepatocellular carcinoma. *J. Proteome Res.* **2008**, *7* (10), 4289–98.
- (14) Sun, Y.; Mi, W.; Cai, J.; Ying, W.; Liu, F.; Lu, H.; Qiao, Y.; Jia, W.; Bi, X.; Lu, N.; Liu, S.; Qian, X.; Zhao, X. Quantitative proteomic signature of liver cancer cells: tissue transglutaminase 2 could be a novel protein candidate of human hepatocellular carcinoma. *J. Proteome Res.* **2008**, *7* (9), 3847–59.
- (15) Shi, R.; Kumar, C.; Zougman, A.; Zhang, Y.; Podtelejnikov, A.; Cox, J.; Wisniewski, J. R.; Mann, M. Analysis of the mouse liver proteome using advanced mass spectrometry. *J. Proteome Res.* **2007**, *6* (8), 2963–72.
- (16) Wisniewski, J. R.; Zougman, A.; Nagaraj, N.; Mann, M. Universal sample preparation method for proteome analysis. *Nat. Methods* **2009**, *6* (5), 359–62.
- (17) Chen, E. I.; Hewel, J.; Felding-Habermann, B.; Yates, J. R. 3rd Large scale protein profiling by combination of protein fractionation and multidimensional protein identification technology (MudPIT). *Mol. Cell. Proteomics* **2006**, *5* (1), 53–6.
- (18) Wang, H.; Clouthier, S. G.; Galchev, V.; Misek, D. E.; Duffner, U.; Min, C. K.; Zhao, R.; Tra, J.; Omenn, G. S.; Ferrara, J. L.; Hanash, S. M. Intact-protein-based high-resolution three-dimensional quantitative analysis system for proteome profiling of biological fluids. *Mol. Cell. Proteomics* **2005**, *4* (5), 618–25.

- (19) Sun, A.; Jiang, Y.; Wang, X.; Liu, Q.; Zhong, F.; He, Q.; Guan, W.; Li, H.; Sun, Y.; Shi, L.; Yu, H.; Yang, D.; Xu, Y.; Song, Y.; Tong, W.; Li, D.; Lin, C.; Hao, Y.; Geng, C.; Yun, D.; Zhang, X.; Yuan, X.; Chen, P.; Zhu, Y.; Li, Y.; Liang, S.; Zhao, X.; Liu, S.; He, F. Liverbase: a comprehensive view of human liver biology. *J. Proteome Res.* **2010**, *9* (1), 50–8.
- (20) Du, R.; Long, J.; Yao, J.; Dong, Y.; Yang, X.; Tang, S.; Zuo, S.; He, Y.; Chen, X. Subcellular quantitative proteomics reveals multiple pathway cross-talk that coordinates specific signaling and transcriptional regulation for the early host response to LPS. *J. Proteome Res.* **2010**, *9* (4), 1805–21.
- (21) Yuan, B. Z.; Jefferson, A. M.; Millecchia, L.; Popescu, N. C.; Reynolds, S. H. Morphological changes and nuclear translocation of DLC1 tumor suppressor protein precede apoptosis in human non-small cell lung carcinoma cells. *Exp. Cell Res.* **2007**, *313* (18), 3868–80.
- (22) McKinney, K. Q.; Lee, Y. Y.; Choi, H. S.; Groseclose, G.; Iannitti, D. A.; Martinie, J. B.; Russo, M. W.; Lundgren, D. H.; Han, D. K.; Bonkovsky, H. L.; Hwang, S. I. Discovery of putative pancreatic cancer biomarkers using subcellular proteomics. *J. Proteomics* **2011**, *74* (1), 79–88.
- (23) Liu, H.; Sadygov, R. G.; Yates, J. R., 3rd A model for random sampling and estimation of relative protein abundance in shotgun proteomics. *Anal. Chem.* **2004**, *76* (14), 4193–201.
- (24) Pavelka, N.; Pelizzola, M.; Vizzardelli, C.; Capozzoli, M.; Splendiani, A.; Granucci, F.; Ricciardi-Castagnoli, P. A power law global error model for the identification of differentially expressed genes in microarray data. *BMC Bioinform.* **2004**, *5*, 203.
- (25) Pavelka, N.; Fournier, M. L.; Swanson, S. K.; Pelizzola, M.; Ricciardi-Castagnoli, P.; Florens, L.; Washburn, M. P. Statistical similarities between transcriptomics and quantitative shotgun proteomics data. *Mol. Cell. Proteomics* **2008**, *7* (4), 631–44.
- (26) Hou, W.; Tian, Q.; Zheng, J.; Bonkovsky, H. L. MicroRNA-196 represses Bach1 protein and hepatitis C virus gene expression in human hepatoma cells expressing hepatitis C viral proteins. *Hepatology* **2011**, *51* (5), 1494–504.
- (27) Nesvizhskii, A. I.; Keller, A.; Kolker, E.; Aebersold, R. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* **2003**, *75* (17), 4646–58.
- (28) Peng, J.; Elias, J. E.; Thoreen, C. C.; Licklider, L. J.; Gygi, S. P. Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC–MS/MS) for large-scale protein analysis: the yeast proteome. *J. Proteome Res.* **2003**, *2* (1), 43–50.
- (29) Lundgren, D. H.; Hwang, S. I.; Wu, L.; Han, D. K. Role of spectral counting in quantitative proteomics. *Expert Rev. Proteomics* **2010**, *7* (1), 39–53.
- (30) Benjamini, Y.; Hochberg, Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J. R. Stat. Soc., Ser. B* **1995**, *57* (1), 289–300.
- (31) Han, M. H.; Hwang, S. I.; Roy, D. B.; Lundgren, D. H.; Price, J. V.; Ousman, S. S.; Fernald, G. H.; Gerlitz, B.; Robinson, W. H.; Baranzini, S. E.; Grinnell, B. W.; Raine, C. S.; Sobel, R. A.; Han, D. K.; Steinman, L. Proteomic analysis of active multiple sclerosis lesions reveals therapeutic targets. *Nature* **2008**, *451* (7182), 1076–81.
- (32) Omenn, G. S.; States, D. J.; Adamski, M.; Blackwell, T. W.; Menon, R.; Hermjakob, H.; Apweiler, R.; Haab, B. B.; Simpson, R. J.; Eddes, J. S.; Kapp, E. A.; Moritz, R. L.; Chan, D. W.; Rai, A. J.; Admon, A.; Aebersold, R.; Eng, J.; Hancock, W. S.; Hefta, S. A.; Meyer, H.; Paik, Y. K.; Yoo, J. S.; Ping, P.; Pounds, J.; Adkins, J.; Qian, X.; Wang, R.; Wasinger, V.; Wu, C. Y.; Zhao, X.; Zeng, R.; Archakov, A.; Tsugita, A.; Beer, I.; Pandey, A.; Pisano, M.; Andrews, P.; Tammen, H.; Speicher, D. W.; Hanash, S. M. Overview of the HUPO Plasma Proteome Project: results from the pilot phase with 35 collaborating laboratories and multiple analytical groups, generating a core dataset of 3020 proteins and a publicly-available database. *Proteomics* **2005**, *5* (13), 3226–45.
- (33) Hanash, S. M.; Pitteri, S. J.; Faca, V. M. Mining the plasma proteome for cancer biomarkers. *Nature* **2008**, *452* (7187), 571–9.
- (34) De Giorgi, V.; Monaco, A.; Worchech, A.; Tornesello, M.; Izzo, F.; Buonaguro, L.; Marincola, F. M.; Wang, E.; Buonaguro, F. M. Gene profiling, biomarkers and pathways characterizing HCV-related hepatocellular carcinoma. *J. Transl. Med.* **2009**, *7*, 85.
- (35) Furge, K. A.; Dykema, K. J.; Ho, C.; Chen, X. Comparison of array-based comparative genomic hybridization with gene expression-based regional expression biases to identify genetic abnormalities in hepatocellular carcinoma. *BMC Genomics* **2005**, *6* (1), 67.
- (36) Speicher, K. D.; Kolbas, O.; Harper, S.; Speicher, D. W. Systematic analysis of peptide recoveries from in-gel digestions for protein identifications in proteome studies. *J. Biomol. Tech.* **2000**, *11* (2), 74–86.
- (37) Shevchenko, A.; Wilm, M.; Vorm, O.; Mann, M. Mass spectrometric sequencing of proteins silver-stained polyacrylamide gels. *Anal. Chem.* **1996**, *68* (5), 850–8.
- (38) Shevchenko, A.; Tomas, H.; Havlis, J.; Olsen, J. V.; Mann, M. In-gel digestion for mass spectrometric characterization of proteins and proteomes. *Nat. Protoc.* **2006**, *1* (6), 2856–60.
- (39) Shevchenko, A.; Sunyaev, S.; Liska, A.; Bork, P. Nanoelectrospray tandem mass spectrometry and sequence similarity searching for identification of proteins from organisms with unknown genomes. *Methods Mol. Biol.* **2003**, *211*, 221–34.
- (40) Zhang, Z.; Carmichael, G. G. The fate of dsRNA in the nucleus: a p54(nrb)-containing complex mediates the nuclear retention of promiscuously A-to-I edited RNAs. *Cell* **2001**, *106* (4), 465–75.
- (41) Senderek, J.; Garvey, S. M.; Krieger, M.; Guergueltcheva, V.; Urtizberea, A.; Roos, A.; Elbracht, M.; Stendel, C.; Tournev, I.; Mihailova, V.; Feit, H.; Tramonte, J.; Hedera, P.; Crooks, K.; Bergmann, C.; Rudnik-Schoneborn, S.; Zerres, K.; Lochmuller, H.; Seboun, E.; Weis, J.; Beckmann, J. S.; Hauser, M. A.; Jackson, C. E. Autosomal-dominant distal myopathy associated with a recurrent missense mutation in the gene encoding the nuclear matrix protein, matrin 3. *Am. J. Hum. Genet.* **2009**, *84* (4), 511–8.
- (42) Roy, L.; Laboissiere, S.; Abdou, E.; Thibault, G.; Hamel, N.; Taheri, M.; Boismenu, D.; Lanoix, J.; Kearney, R. E.; Paiement, J. Proteomic analysis of the transitional endoplasmic reticulum in hepatocellular carcinoma: an organelle perspective on cancer. *Biochim. Biophys. Acta* **2010**, *1804* (9), 1869–81.
- (43) Ende, S.; Fuhr, M.; Pak, S. J.; Zabel, B. U.; Winterpacht, A. LETM1, a novel gene encoding a putative EF-hand Ca(2+)-binding protein, flanks the Wolf-Hirschhorn syndrome (WHS) critical region and is deleted in most WHS patients. *Genomics* **1999**, *60* (2), 218–25.
- (44) Tamai, S.; Iida, H.; Yokota, S.; Sayano, T.; Kiguchiya, S.; Ishihara, N.; Hayashi, J.; Mihara, K.; Oka, T. Characterization of the mitochondrial protein LETM1, which maintains the mitochondrial tubular shapes and interacts with the AAA-ATPase BCS1L. *J. Cell Sci.* **2008**, *121* (Pt 15), 2588–600.
- (45) Schlickum, S.; Moghekar, A.; Simpson, J. C.; Steglich, C.; O'Brien, R. J.; Winterpacht, A.; Ende, S. U. LETM1, a gene deleted in Wolf-Hirschhorn syndrome, encodes an evolutionarily conserved mitochondrial protein. *Genomics* **2004**, *83* (2), 254–61.
- (46) Nowikovsky, K.; Froschauer, E. M.; Zsurka, G.; Samaj, J.; Reipert, S.; Kolisek, M.; Wiesenbergerger, G.; Schwelyn, R. J. The LETM1/YOL027 gene family encodes a factor of the mitochondrial K+ homeostasis with a potential role in the Wolf-Hirschhorn syndrome. *J. Biol. Chem.* **2004**, *279* (29), 30307–15.
- (47) Brill, S.; Li, S.; Lyman, C. W.; Church, D. M.; Wasmuth, J. J.; Weissbach, L.; Bernards, A.; Snijders, A. J. The Ras GTPase-activating-protein-related human protein IQGAP2 harbors a potential actin binding domain and interacts with calmodulin and Rho family GTPases. *Mol. Cell. Biol.* **1996**, *16* (9), 4869–78.
- (48) Schmidt, V. A.; Chiariello, C. S.; Capilla, E.; Miller, F.; Bahou, W. F. Development of hepatocellular carcinoma in Iggap2-deficient mice is IQGAP1 dependent. *Mol. Cell. Biol.* **2008**, *28* (5), 1489–502.