

Shotgun Proteomics Analysis of Differentially Expressed Urinary Proteins Involved in the Hepatocellular Carcinoma

Chao-Jen Kuo¹, Shen-Nien Wang^{2#}, Shih-Shin Liang^{3,4#}, Shu-Wen Chi¹, Zong Jing Yu³, Edward Hsi⁵, King-Teh Lee^{2*} and Shyh-Horng Chiou^{1*}

¹Graduate Institute of Medicine, College of Medicine, Kaohsiung Medical University, Taiwan

²Division of Hepatobiliary Surgery, Department of Surgery, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan

³Department of Biotechnology, College of Life Science, Kaohsiung Medical University, Taiwan

⁴Center for Research Resources and Development, Kaohsiung Medical University, Taiwan

⁵Department of Medical Research, Kaohsiung Medical University Hospital, Taiwan

#These authors contributed equally to this work

Abstract

Numerous investigations underlying the hepatocellular carcinoma (HCC) diagnosis as well as detection at an early stage based on hepatitis B virus surface antigen (HbsAg) concentration in serum and aflatoxin metabolites in urine have been commonly reported in the literature. To date, these biomarkers, even though specific and accurate, are not universal for the detection of HCC elicited by all causative factors. In addition, potential biomarkers may be present at low concentrations in contrast to the presence of abundant interfering proteins with a wide dynamic range. The aim of this study is to establish an operational analysis platform of effective and noninvasive diagnostic tool with a high sensitivity to explore protein expression profiles by shotgun proteomics using nano-liquid chromatography coupled tandem mass spectrometry (nanoLC-MS/MS) and stable isotope dimethyl labeling. Differentially expressed urinary proteins were identified and compared by the mass spectral patterns of their peptide fragments generated from protease digestion. In our results, the quantitative proteomic analysis of the differentially expressed proteins in urine identified at least 21 protein biomarker candidates with high confidence levels. We have further identified 14 proteins with up-regulation (stable isotope D/H ratio ≥ 1.5) and 7 with down-regulation (D/H ratio ≤ 0.6). The systematic decrease or increase of these identified marker proteins may potentially reflect the morphological aberrations and diseased stages of liver throughout progressive developments of HCC. The results would place a firm foundation for future work regarding validation and clinical translation of some identified biomarkers into targeted diagnosis and therapy for various classes of HCC.

Keywords: Quantitative proteomics; Hepatocellular carcinoma (HCC); Shotgun proteomics analysis; Nano-liquid chromatography coupled tandem mass spectrometry (nanoLC-MS/MS); Stable isotope dimethyl labeling

Introduction

Hepatocellular carcinoma (HCC), being ranked as the seventh most common malignant tumor in women and the fifth in men worldwide [1], is the first leading cause of cancer death in epidemiological studies of Taiwan in recent decades [2]. It has also been noted that HCC was prevalent in middle and western Africa, and eastern and south-eastern Asia; furthermore, over 80% of HCC incidences occur in developing countries [1]. Many etiological factors accountable for a wide spectrum of clinical manifestations were found to be associated generally with chronic infection with hepatitis B or C virus (HBV, HCV) and cirrhosis [3]. Moreover, several environmental factors including alcoholism, tobacco smoking and dietary exposure to aflatoxins could also account partly for a high incidence of HCC [4,5]. Many patients detected with HCC were found to have cirrhosis when they were initially diagnosed as sufferers of chronic liver disease lacking long-term clinical care and appropriate treatment. This abnormal cirrhosis condition results in an increase in the replacement of normal tissue with fibrous tissue leading to the loss of functional liver cells and the development of HCC [6].

The mortality rate for HCC exceeds 30 cases per 10,000 population and most cases are resistant to traditional chemotherapy and radiotherapy [7]. A wide variety of chemotherapeutic agents currently in use include fluorouracil, doxorubicin, mitoxantrone, cisplatin, mitomycin C, epirubicin, interferon-alpha and tegafur. However, no curative regimen has been found to date. The drug response and prolongation of survival are usually minimal (a few months or less),

and there is a significant morbidity associated with poor treatment [8]. Curative surgery of HCC is feasible for only about 30% of patients. Transarterial embolization or chemoembolization (TAE/TACE) has been demonstrated to provide some survival benefits if tumors are confined to the localized area of liver and no evidence of portal vein thrombosis is found [9]. Therefore, in general diagnosis or detection at an early stage of HCC is considered to be essential to allow favorable and positive clinical treatments for increasing the life expectancy of HCC patients. Some screening tools such as the measurement of serum alpha-fetoprotein (AFP) or an abdominal ultrasonography at regular intervals targeting high risk populations were also applied to the detection of HCC at an early stage. Unfortunately, poor sensitivity and specificity of AFP and the need of an operator's expertise required on the ultrasonographic evaluation limited their prognostic use.

***Corresponding authors:** Shyh-Horng Chiou, Quantitative Proteomics Center and Graduate Institute of Medicine, Kaohsiung Medical University, Kaohsiung, Taiwan, Tel: +886-7-3220377; Fax: +886-7-3133434; E-mail: shchiou@kmu.edu.tw

King-Teh Lee, Division of Hepatobiliary Surgery, Department of Surgery, Kaohsiung Medical University Hospital, Kaohsiung Medical University, Kaohsiung, Taiwan, Tel: +886-7-3121101(ext.7651); Fax: +886-7-3216992; E-mail: ktlee@kmu.edu.tw

Received January 06, 2014; **Accepted** February 01, 2014; **Published** February 05, 2014

Citation: Kuo CJ, Wang SN, Liang SS, Chi SW, Yu ZJ, et al. (2014) Shotgun Proteomics Analysis of Differentially Expressed Urinary Proteins Involved in the Hepatocellular Carcinoma. J Proteomics Bioinform 7: 034-040. doi:10.4172/0974-276X.1000300

Copyright: © 2014 Kuo CJ, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Combined use of AFP testing and ultrasonography was also reported to increase false-positive rates [10]. Other tests, including *Lens culinaris* agglutinin-reactive AFP and des-gamma carboxyprothrombin (DCP), are currently being evaluated and validated [11]. Owing to the lack of accurate and specific biomarkers for the assessment of HCC incidence at an early stage, HCC is thus considered to be difficult for detection and usually diagnosed as an incurable disease when diagnosed at a late diseased stage. Therefore it is very urgent and imperative to discover biomarkers with prognosis potential coupled with the development of curative therapy in the effective management of HCC.

Although greater emphasis in biological research is being directed toward a comprehensive global analysis of protein expression profiles for biomarker discovery, reliable and high-throughput proteome-wide comparative analyses for some diseases has not existed until the advent of current proteomics instrumentation. The qualitative and quantitative studies of proteins by means of fast-evolving and state-of-the-art proteomics methodologies have provided a firm basis for understanding the complex proteome profiles of total protein mixtures from various sample sources such as tissues, cells, plasma and urine [12,13]. A major phase of various proteomics strategies lies in the determination of protein identity (Protein ID) of interest using analytical "fingerprints" or peptide mass fingerprinting (PMF) generated by digestion of proteins with cleavage-specific enzymes such as trypsin or some other well-characterized proteases, from which tandem mass (MS/MS) spectra of peptide fragments can then be used for comparison and confirmation of protein ID in available sequence databanks. The strategy based on "bottom-up" proteomic approach by means of the direct analysis of peptides generated from protein digestion by high-resolution liquid chromatographies coupled with tandem MS/MS spectrometry (nanoLC-MS/MS) has facilitated the so-called "shotgun proteomics" for the identification of protein mixtures from any tissues of interest. Basically shotgun proteomics detects PMF profiles throughout the whole cellular proteome based on a random statistical sampling method similar to that of shotgun genomics in the decoding of human genome sequences. Various MS/MS spectra can be algorithmically compared with predicted peptide spectra from sequence databases to identify the respective proteins. This shotgun proteomics approach is capable of characterizing proteins directly from entire tissue or cell lysates [14-16].

In this study, we aim to establish a high-throughput operation platform of effective and noninvasive diagnostic tool for early detection of HCC biomarkers. To attain this goal, we have made an effort to characterize and compare the urinary proteins between diseased and control groups in order to identify potential biomarker candidates by means of gel-free shotgun proteomic analysis coupled with stable isotope dimethyl labeling [12,16,17] and nanoLC-MS/MS [17-19]. The proteomic analysis at the global cellular level reported herein will lay a firm foundation for future work regarding validation and clinical translation of some identified biomarkers into targeted diagnosis and therapy for HCC.

Materials and Methods

Trichloroacetic acid (TCA), trifluoroacetic acid (TFA), dithiothreitol (DTT), iodoacetamide (IAM), ethylenediaminetetraacetic acid (EDTA), sodium deoxycholate, sodium fluoride (NaF), formaldehyde- H_2 , formaldehyde- D_2 and ammonium bicarbonate (NH_4HCO_3), Triton X-100 were purchased from Sigma Aldrich (St. Louis, MO). Acetonitrile (ACN) and sodium phosphate were obtained from Merck

(Darmstadt, Germany). Formic acid (FA), sodium acetate, sodium cyanoborohydride and sodium chloride (NaCl) were purchased from Riedel-de Haven (Seelze, Germany). Protease inhibitors (Complete™ Mini) were purchased from Roche (Mannheim, Germany). Sodium dodecyl sulfate (SDS) and urea were purchased from Amresco (Solon, OH). Modified sequencing-grade trypsin for in-gel digestion was purchased from Promega (Madison, WI). Quantitative reagent for protein contents was purchased from Bio-Rad (Hercules, CA). Water was deionized to 18 M Ω by a Milli-Q system (Millipore, Bedford, MA).

Sample collection

All the procedures used in this study were approved by the ethical committee of clinical research at Kaohsiung Medical University Hospital. We collected urine from patients who were diagnosed as cases of HCC incidence and never underwent cholecystectomy as the disease group. We also got agreement from patients diagnosed with non-HCC incidence but underwent cholecystectomy for urine collection as the normal control. 50 mL urine for each individual were harvested, concentrated by centrifugation and assayed for determination of total protein concentration using Coomassie protein assay reagent, and subsequently were stored at -80°C until being analyzed.

Dimethyl labeling and peptide preparation

Volumes of urine containing 100 μ g of total proteins were adjusted to 60 μ L and treated with 0.7 μ L of 1 DTT and 9.3 μ L of 7.5% SDS at 95°C for 5 min before reduction. After the reaction, lysates were further treated with 8 μ L of 50 mM IAM at room temperature for alkylation in the dark for 30 min; subsequently proteins were precipitated by adding 52 μ L of 50% TCA and incubated on ice for 15 min. After removing the supernatant by centrifugation at 13,000 x g for 5 min, the collected proteins were washed with 150 μ L of 10% TCA, vortexed and centrifuged at 13,000 x g for 10 min. The precipitated proteins were washed again with 250 μ L distilled H_2O , vortexed and centrifuged thrice under the same condition. The resultant pellets were resuspended with 50 mM NH_4HCO_3 (pH 8.5), then digested with 4 μ g of trypsin for 8 h at 37°C and further dried in a vacuum centrifuge to remove NH_4HCO_3 . The lyophilized peptides from HCC and normal urine re-dissolved in 180 μ L of 100 mM sodium acetate at pH 5.5 were treated with 20 μ L of 4% formaldehyde- H_2 and 20 μ L 4% formaldehyde- D_2 , respectively [17-19] and mixed thoroughly. The mixtures were vortexed for 5 min, immediately followed by the addition of 10 μ L of 0.6 M sodium cyanoborohydride and vortexed for 1 h at room temperature. The resultant liquids were acidified by 10% TFA/ H_2O to pH 2.0~3.0 and applied onto the in-house reverse-phase C18 column pre-equilibrated with 200 μ L of 0.1% TFA/ H_2O (pH 2.0~3.0) for desalting. The column was also washed with 200 μ L of 0.1% TFA/ H_2O (pH 3.0) and then eluted with a stepwise ACN gradient from 50% to 100% in 0.1% TFA at room temperature.

Hydrophilic interaction chromatography (HILIC) for peptide separation

HILIC was performed on an L-7100 pump system with quaternary gradient capability (Hitachi, Tokyo, Japan) using a TSK gel Amide-80 HILIC column (2.0 \times 150 mm, 3 μ m; Tosoh Biosciences, Tokyo, Japan) [20-22] with a flow rate of 200 μ L/min. Two buffers were used for gradient elution: solvent (A), 0.1% TFA in water, and solvent (B), 0.1% TFA in 100% ACN. The eluted fractions after being desalted from the in-house reverse-phase C18 column were each dissolved in 25 μ L of solution containing 85% ACN and 0.1% TFA and then injected into the

20 μ L sample loop. The gradient was processed as follows: 98% (B) for 5 min, 98-85% (B) for 5 min, 85-0% (B) for 40 min, 0% (B) for 5 min, 0-98% (B) for 2 min and 98% (B) for 3 min. A total of 10 fractions were collected (1.2 mL for each fraction) and dried in a vacuum centrifuge.

NanoLC-MS/MS analysis

The lyophilized powders were reconstituted in 10 μ L of 0.1% FA in H₂O and analyzed by LTQ Orbitrap XL (Thermo Fisher Scientific, San Jose, CA). Reverse-phase nano LC separation was performed on an Agilent 1200 series nanoflow system (Agilent Technologies, Santa Clara, CA). A total of 10 μ L sample from collected fractions was loaded onto an Agilent Zorbax XDB C18 precolumn (0.35 mm, 5 μ m), followed by separation using in-house C18 column (i.d. 75 μ m \times 15 cm, 3 μ m). The mobile phases used were (A) 0.1% FA in water and (B) 0.1% FA in 100% ACN. A linear gradient from 5% to 95% of (B) over a 70-min period at a flow rate of 300 nL/min was applied. The peptides were analyzed in the positive ion mode by applying a voltage of 1.8 kV to the injection needle. The MS was operated in a data-dependent mode, in which one full scan with m/z 400-1600 in the Orbitrap using a scan rate of 30 ms/scan. The fragmentation was performed using the CID mode with collision energy of 35 V. A repeat duration of 30 s was applied to exclude the same m/z ions from the reselection for fragmentation. The Xcalibur software (version 2.0.7, Thermo Fisher Scientific, San Jose, CA) was used for the management of instrument control, data acquisition, and data processing.

Protein database search and characterization

Peptides were identified by peak lists converted from the nanoLC-MS/MS spectra by bioinformatics searching against *Homo sapiens* taxonomy in the Swiss-Prot databases for exact matches using the Mascot search program (<http://www.matrixscience.com>) [23,24]. Parameters were set as follows: a mass tolerance of 10 ppm for precursor ions and 0.8 Da for fragment ions; no missed cleavage site allowed for trypsin; carbamidomethyl cysteine as fixed modification; dimethylation specified as standard of the quantification; oxidized methionine and deamidated asparagine/glutamine as optional modification. Peptides were considered positively identified if their Mascot individual ion score was higher than 20 ($p < 0.05$).

Subsequently, the analysis of peptide quantification ratio (D/H) for normal (hydrogen labeling) and HCC (deuterium labeling) from urine was carried out by Mascot Distiller program (version 2.3, Matrix Science Ltd., London, U.K.) using the average area of the first 3 isotopic peaks across the elution profile. The Mascot search data as well as quantification resulting from each fraction were also merged by this program that combined the peptide ratios matching the same sequence obtained from different fractions or at different retention time and charge state [17]. The identified proteins with up- and down-regulation were further categorized based on their biological process and molecular function using the PANTHER classification system (<http://www.pantherdb.org>) as described in the previous studies [25-27].

Construction of signaling pathways and network analysis of protein interaction

The software program (www.ingenuity.com) from Ingenuity Pathways Analysis (IPA, Ingenuity Systems, Redwood City, CA) was used for deriving the pathways and networks of protein interaction. Protein factors characterized by proteomic analysis were analyzed for

their association with mapping related to canonical pathways deposited in the IPA library.

Results and Discussion

Protein expression levels analyzed by nanoLC-MS/MS

Quantitative proteome analysis by shotgun approach coupled with stable isotope dimethyl labeling has been used in identifying candidate biomarkers or target factors in different types of samples on account of the fact that this approach can detect differentially released proteins at relatively low abundance [28-31]. In this study, we conducted a comparative proteomics investigation of urine between HCC patients and control group by a bottom-up shotgun proteomic approach. A schematic representation of sample processing, separation and the subsequent workflow concerning trypsin digestion, dimethyl labeling and shotgun analysis is shown in Figure 1. Initially, 100 μ g each of total urine proteins from HCC patients and control group was subjected to trypsin digestion and dimethyl labeling. Respective tryptic peptide samples were mixed in a 1:1 (w/w) ratio and then enriched by the reverse-phase C18 column. Owing to the fact that the enriched peptide population was too complex to be fully detected and characterized by a single LC-MS/MS run, the enriched peptides were fractionated by HILIC based on polarity difference, and then harvested into 10 fractions. Each fraction was analyzed by LC-LTQ-Orbitrap and the parameter used in searching for peptide identification was adjusted to allow for no missed cleavage. Most of the peptides were separated from a single or two adjacent HILIC fractions, and peptides identified by the Mascot search program (<http://www.matrixscience.com>) [23,24] were accepted if their individual ion score was higher than 20, which had been a cutoff point used for the lower-quality MS/MS spectra [32-34] Supplementary information.

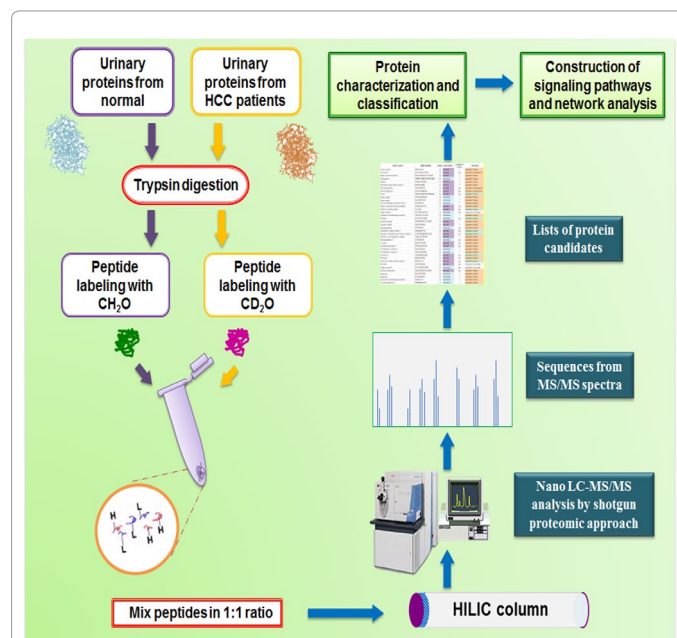


Figure 1: Experimental scheme of the procedures used for the screening of differentially expressed proteins. After enzymatic digestion, peptides were differentially stable isotope dimethyl-labeled and combined prior to desalting and fractionation. The quantitative shotgun analysis of proteome changes from clinical urine samples of normal and HCC patients was carried out by using HILIC-C18 peptide separation and nanoLC-MS/MS coupled with stable isotope dimethyl labeling.

Quantification of identified proteins with differential expression

Once the differentially released proteins with confident identification based on dimethyl labeling, enzyme digestion and peptide mass fingerprinting (PMF) were completed, the peptide quantification ratio (D/H) was obtained by Mascot Distiller program using the average area of the first 3 isotopic peaks across each elution profile [17,18,35]. Data from the publicly available Mascot searching engine as well as quantification results from each fraction were also merged by Mascot Distiller program into one file that combined peptides with more than one peptide matching the same sequence, which were harvested from different fractions or at different retention time and charge state. Herein, the 14 up-regulated (D/H ratio ≥ 1.5) and 7 down-regulated (D/H ratio ≤ 0.6) proteins displayed in at least four of seven comparative urine samples between HCC patients and control group were identified and listed in Table 1. The representative peptide peaks including protein AMBP, prostaglandin-H₂.

D-isomerase (PTGDS), serotransferrin (TF) and alpha-1-antitrypsin (SERPINA1) are shown in Figure 2-5. On the one hand, the isotopic pair (quantification ratio) of nano-LC and MS/MS CID spectra shown in Figure 2 demonstrate that D₄ and H₄-labeled peptides (VVAQGVGIPEDSIFTMADR) derived from protein AMBP, exhibiting m/z values of 1027.04 (+2) and 1025.02 (+2), respectively and D/H ratio of 10.32, were eluted at the same time (~19.19 min) in the HILIC fraction 5. D₄ and H₄-labeled peptides (TMLLQPAGSLGYSYR) from PTGDS, showing m/z values of 896.47 (+2) and 894.45 (+2), respectively and D/H ratio of 5.36, were eluted at the same time (~18.27 min) in the HILIC fraction 5, as shown on the

isotopic pair (quantification ratio) of nano LC and MS/MS CID spectra of Figure 3. These two proteins show increased release among patients based on the quantitative ratios. On the other hand, the isotopic pair of nano-LC and MS/MS CID spectra shown in Figure 4 demonstrate that D₄ and H₄-labeled peptides (DGAGDVAFVK) derived from TF, exhibiting m/z values of 521.78 (+2) and 517.78 (+2), respectively and D/H ratio of 0.05, were eluted at the same time (~15.82 min) in the HILIC fraction 5. D₄ and H₄-labeled peptides (LSITGTYDLK) from SERPINA1, showing m/z values of 587.86 (+2) and 583.84 (+2), respectively and D/H ratio of 0.27, were eluted at the same time (~18.18 min) in the HILIC fraction 4, as shown on the isotopic pair (quantification ratio) of nano-LC and MS/MS CID spectra of Figure 5. Therefore, these two proteins show reduced release among patients. The results from the above data further indicate that no isotopic effect was observed in the two-dimensional HILIC-C18 separation, revealing the excellent separation efficiency of dimethylated peptides as fractionated by HILIC column. The high orthogonality of HILIC and reverse-phase C18 column may have made a great contribution to this efficiency.

All these differentially expressed proteins were further categorized using the PANTHER classification system [25,27,36]. Functional distributions of these identified proteins were shown in Figure 6. It is worthy to note that binding proteins and proteins with catalytic activity account for 28% and 24% of increased release proteins, respectively (Figure 6A) and proteins of receptor type occupied as high as 37% of reduced release proteins (Figure 6B) in the urine of HCC patients. These differentially released proteins were also associated with a variety of biological processes such as cellular process (10%) and metabolic process (21%) shown in Figure 6C; a high proportion of proteins

Protein ID	Gene symbol	D/H ratio	p-value	Number of sample pairs	pl /mass (kDa)	Number of peptides	Swiss prot accession number	Sequence coverage (%)
Increased release								
Protein AMBP	AMBP	2.0 ~ 14.2	0.016*	6/7	5.95 / 40.46	122	P02760	45
Prostaglandin-H ₂ D-isomerase	PTGDS	1.6 ~ 11.8	0.0239*	6/7	7.66 / 21.56	23	P41222	42
Uromodulin	UMOD	2.9 ~ 30.5	0.164	5/7	5.05 / 72.96	90	P07911	44
Beta-2-glycoprotein 1	APOH	2.4 ~ 35.8	0.101	5/7	8.34 / 40.55	2	P02749	14
Plasma serine protease inhibitor	SERPINA5	2.3 ~ 17.3	0.0417*	5/7	9.3 / 46.53	19	P05154	28
Lysosomal alpha-glucosidase	GAA	2.6 ~ 8.5	0.013*	5/7	5.62 / 106.59	5	P10253	17
Ig kappa chain V-III region SIE	KV302	1.6 ~ 18.1	0.216	5/7	8.72 / 11.98	4	P01620	39
Ig lambda-2 chain C	IGLC2	2.1 ~ 10.9	0.157	4/7	6.92 / 11.69	17	P0CG05	69
Kininogen-1	KNG1	2.7 ~ 28.4	0.108	4/7	6.34 / 74.43	62	P01042	37
Basement membrane-specific heparan sulfate proteoglycan core protein	HSPG2	3.2 ~ 13.2	0.082	4/7	6.06 / 481.43	35	Q2VPA1	10
Polymeric immunoglobulin receptor	PIGR	2.0 ~ 17.7	0.165	4/7	5.58 / 85.78	5	P01833	24
Ribonuclease	RNASE1	4.5 ~ 30.3	0.099	4/7	9.1 / 18.38	13	P07998	37
Ig kappa chain C region	IGKC	1.6 ~ 13.7	0.158	4/7	5.58 / 12.00	62	P01834	80
Pro-epidermal growth factor	EGF	1.8 ~ 19.5	0.243	4/7	5.53 / 139.41	2	P01133	12
Reduced release								
Serotransferrin	TF	0.03 ~ 0.14	0.181	6/7	6.81 / 81.15	35	P02787	33
Alpha-1-antitrypsin	SERPINA1	0.08 ~ 0.31	<0.0001*	6/7	5.37 / 47.83	49	P01009	36
Ig gamma-2 chain C	IGHG2	0.03 ~ 0.47	0.920	5/7	7.66 / 37.18	31	P01859	31
Alpha-1B-glycoprotein	A1BG	0.07 ~ 0.35	0.976	5/7	5.56 / 55.10	3	P04217	24
Zinc-alpha-2-glycoprotein	AZGP1	0.05 ~ 0.2	0.500	4/7	5.71 / 35.11	9	P25311	39
Zinc-alpha-2-glycoprotein	SERPINA3	0.28 ~ 0.5	0.716	4/7	5.33 / 48.52	32	P01011	36
Serum albumin	ALB	0.04 ~ 0.41	0.003	4/7	5.92 / 73.00	322	P02768	66

Protein expression ratios of D/H shown here cover the range of seven measurements.

D = Deuterium labeling on peptide fragments of proteins from HCC urine; H = Hydrogen labeling on peptide fragments of proteins from control of normal counterpart)

*Statistically significant data by student t-test (p<0.05)

Table 1: Increased (D/H ratio ≥ 1.5) and decreased expression (D/H ratio ≤ 0.6) proteins were displayed in at least four of seven hepatocellular carcinoma (HCC) urine sample pairs when identified by nanoLC-MS/MS coupled with stable isotope dimethyl labeling.

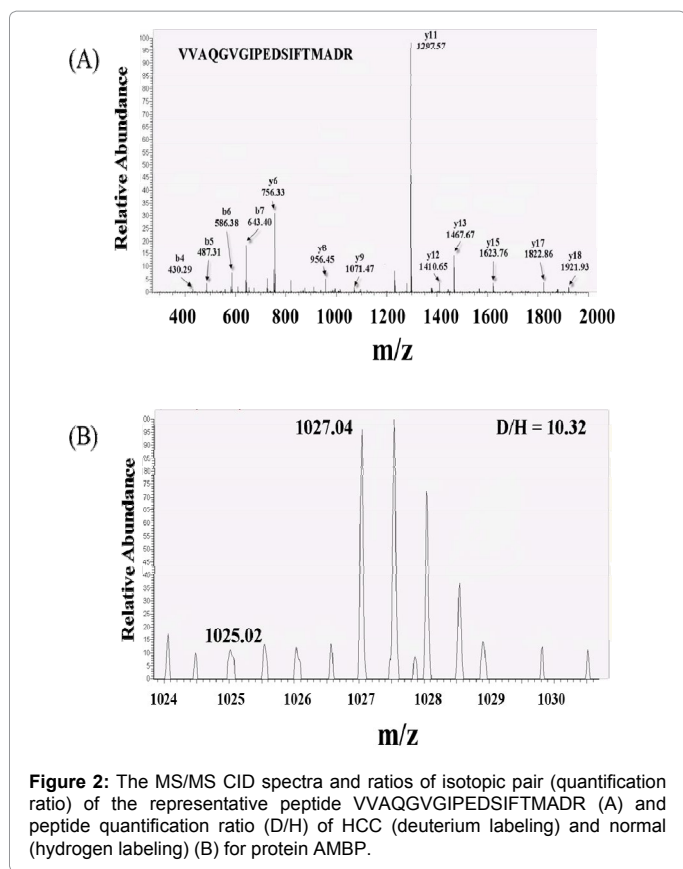


Figure 2: The MS/MS CID spectra and ratios of isotopic pair (quantification ratio) of the representative peptide VVAQGVGIPEDSIFTMADR (A) and peptide quantification ratio (D/H) of HCC (deuterium labeling) and normal (hydrogen labeling) (B) for protein AMBP.

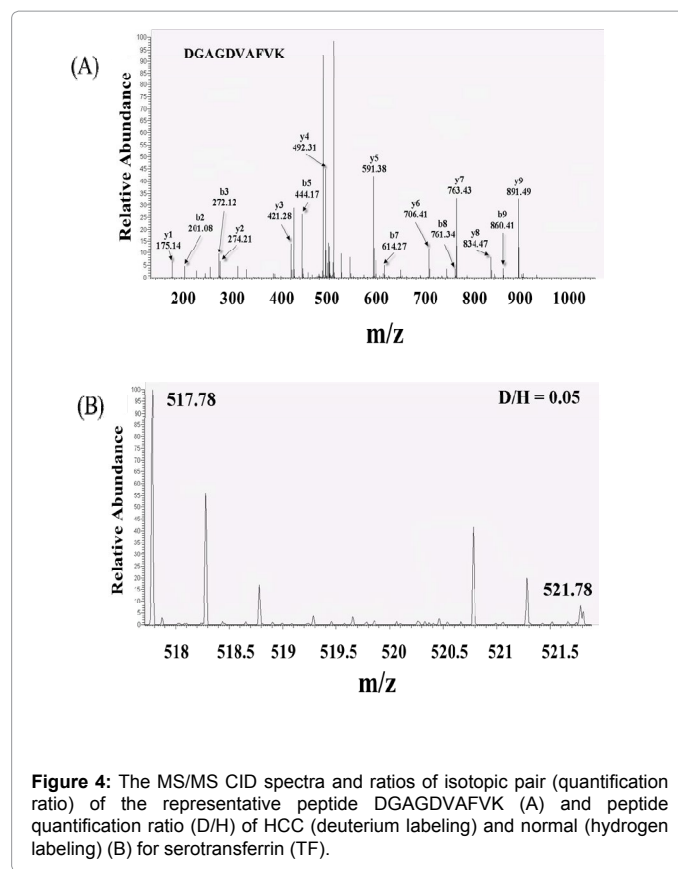


Figure 4: The MS/MS CID spectra and ratios of isotopic pair (quantification ratio) of the representative peptide DGAGDVAFVK (A) and peptide quantification ratio (D/H) of HCC (deuterium labeling) and normal (hydrogen labeling) (B) for serotransferrin (TF).

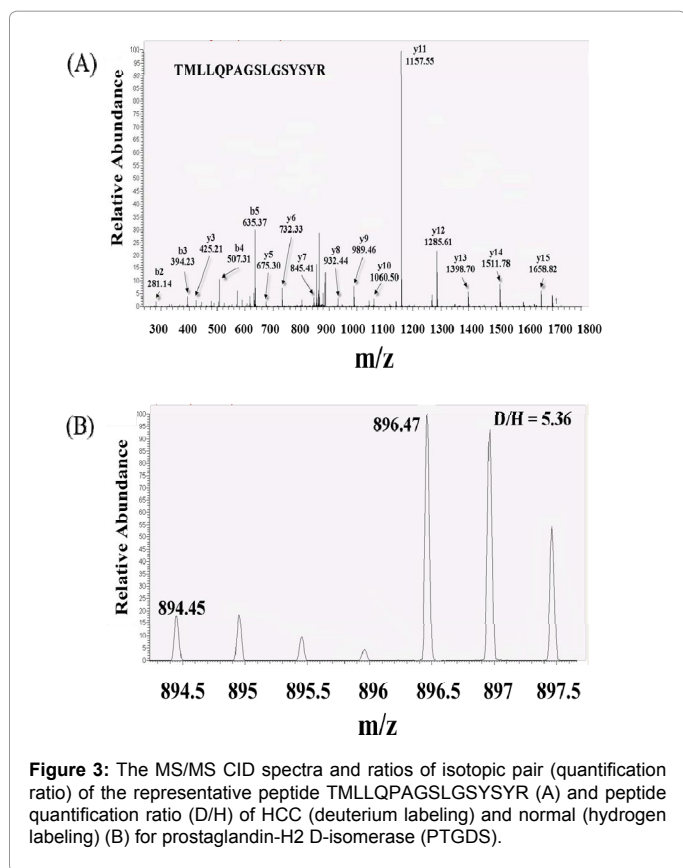


Figure 3: The MS/MS CID spectra and ratios of isotopic pair (quantification ratio) of the representative peptide TMLLQPAGSLGSYSYR (A) and peptide quantification ratio (D/H) of HCC (deuterium labeling) and normal (hydrogen labeling) (B) for prostaglandin-H2 D-isomerase (PTGDS).

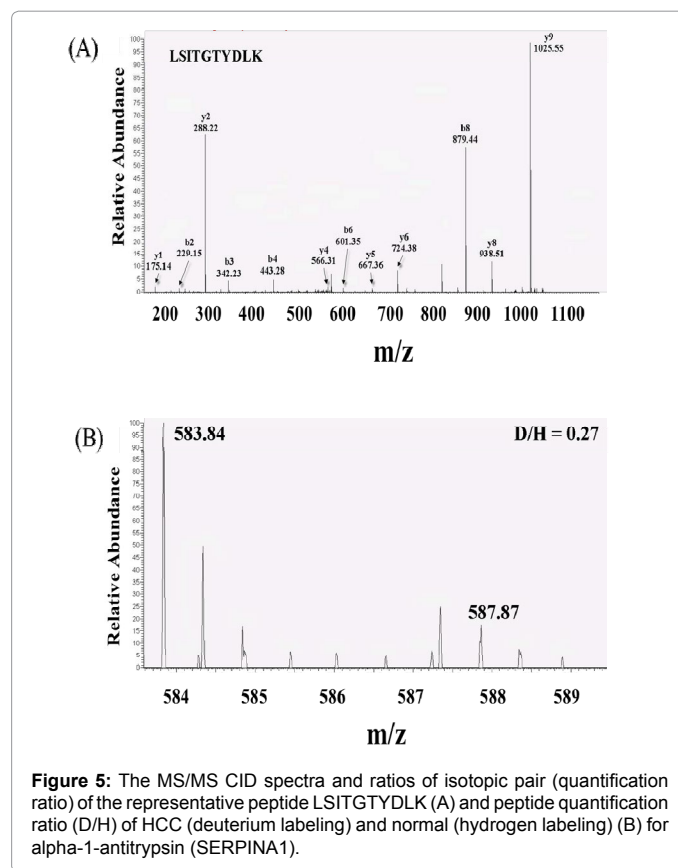


Figure 5: The MS/MS CID spectra and ratios of isotopic pair (quantification ratio) of the representative peptide LSITGTYDLK (A) and peptide quantification ratio (D/H) of HCC (deuterium labeling) and normal (hydrogen labeling) (B) for alpha-1-antitrypsin (SERPINA1).

involved in the process of immune system (12%) showed reduced release, as shown in Figure 6D. Therefore, our current observation demonstrates that decreased and increased expression of these proteins implied the likelihood of their involvement in the pathologic statuses of carcinogenesis and differentiation of HCC cells.

Construction of signaling pathways and network analysis of protein interaction

Using a panel of these identified proteins, we further cluster them into a possible connection network based on the biochemical categorization to put forward a simulation scheme for the prospective signaling pathways governing the maintenance and progression of carcinogenic status in the liver tissue. In Figure 7, these identified

proteins (shown in red color) mapped to canonical pathways derived by the Ingenuity Pathways Analysis (IPA, Ingenuity Systems) databank were displayed with different shapes to signify the disparate and diverse functions. Proteins reported in the literature and canonical pathway database based on their functional annotation were shown in white color and subjected to the association analysis and simulation of possible molecular interaction with our identified proteins. All the gray arrows designate the biological interrelationships between molecules. All arrows in the figure were supported by at least one reference from the literature, textbooks, or canonical information stored in the Ingenuity Knowledge Base. The increased release proteins including pro-epidermal growth factor (EGF), kininogen-1 (KNG1), beta-2-glycoprotein 1 (APOH) and polymeric immunoglobulin receptor (PIGR), to some extent, are involved in inflammatory responses; simultaneously, basement membrane-specific heparan sulfate proteoglycan core protein (HSPG2) and ribonuclease (RNASE1) were categorized to be involved in the proliferation of cancer cells. Namely, HCC was not characterized with a singular and clear-cut enzymatic or cytoskeleton alteration but with a series of complex and diverse functional changes.

Several proteins identified by our shotgun approach were not recruited to the canonical pathways mapped in the database owing to the fact that these proteins were not linked to functional interaction; however, the importance of these unmapped proteins with up-regulated release cannot be overlooked. The availability and suitability of these identified proteins employed as candidate biomarkers will be validated by subjecting them to next phases of verification and validation using ion scanning of peptides measured and quantified in multiple reaction monitoring (MRM) mode of nanoLC-MS/MS analysis. In addition, a much larger number of matched sample pairs should be essential for being able to discriminate the subtle and yet crucial differences of released proteins between HCC and corresponding normal counterparts. After completion of the initial phase of biomarker discovery based on limited sample pairs in this preliminary pilot study, we are currently embarking the second phase of biomarker verification based on the identified candidate marker proteins and an expanded scale of urine sample collection of HCC and some other patients of different diseases [37,38].

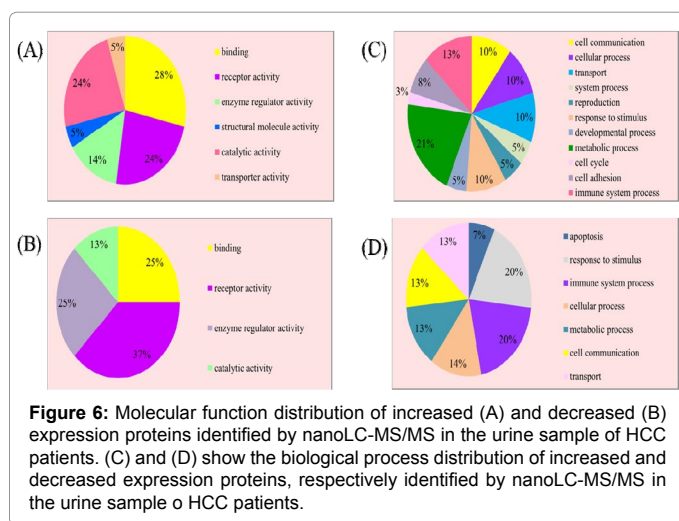


Figure 6: Molecular function distribution of increased (A) and decreased (B) expression proteins identified by nanoLC-MS/MS in the urine sample of HCC patients. (C) and (D) show the biological process distribution of increased and decreased expression proteins, respectively identified by nanoLC-MS/MS in the urine sample of HCC patients.

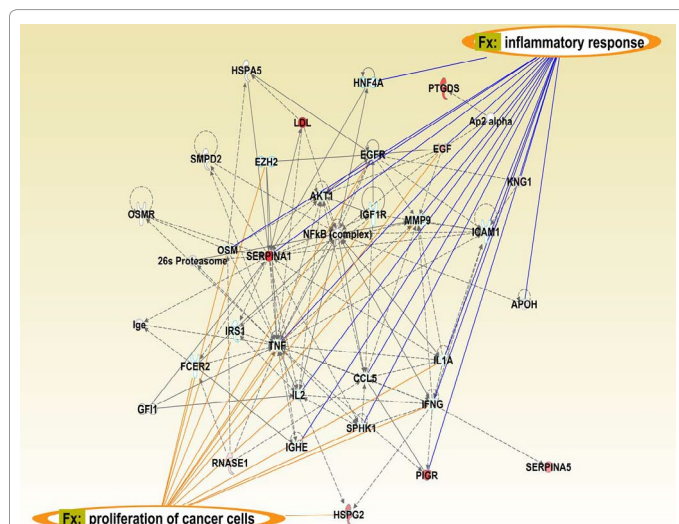


Figure 7: Schematic representation of derived pathways associated with liver cancer and inflammation. The networks of these identified proteins mapping to the canonical pathways from Ingenuity Pathways Analysis (IPA, Ingenuity Systems) library were employed for the analysis of proteins with increased or decreased expression. Identified proteins shown in red color were displayed with different shapes to indicate different functions. The biological interrelationships between molecules were represented as arrows. All drawn arrows were cited and supported by at least one reference from the literature, textbooks, or canonical information stored in the Ingenuity Knowledge Base. Fx denoted for function.

Conclusion

Hepatocellular carcinomas (HCCs) encompass different etiology and pathological manifestations coupled with heterogenic genomic alterations leading to high complexity and intractable therapy and treatment. Collectively, the severity of HCC involves a variety of protein factors which play some regulatory roles in metabolic coordination of physiological functions. The systematic decrease and increase of these proteins may be reflective of the dysfunction of liver cells, followed by morphological aberrations upon progressive developments of HCC. The comparative proteome data from urine samples may help not only offer a novel approach to further understand the mechanism(s) underlying the development and the associated metabolic signaling pathways entailed in liver carcinogenesis, but also develop potential and valuable biomarker candidates useful for the non-invasive diagnosis and prognosis.

Acknowledgements

This work was supported in part by Kaohsiung Medical University (KMU), Academia Sinica and the National Science Council (NSC Grants 99-2745-B-037-005 and 102-2314-B-037-037-MY2 to S.H. Chiou and K.T. Lee), Taipei, Taiwan. We thank the core facility grant from the quantitative proteomics center (NSC Grant

99-2745-B-037-005) at the Center for Research Resources and Development (CRRD), Kaohsiung Medical University under the auspices of National Science Council.

References

1. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, et al. (2010) Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 127: 2893-2917.
2. Kee KM, Lu SN (2011) Hospital- and community-based screenings for hepatocellular carcinoma in Taiwan. *Oncology* 81: 36-40.
3. Montalto G, Cervello M, Giannitrapani L, Dantona F, Terranova A, et al. (2002) Epidemiology, risk factors, and natural history of hepatocellular carcinoma. *Ann N Y Acad Sci* 963: 13-20.
4. McGlynn KA, London WT (2005) Epidemiology and natural history of hepatocellular carcinoma. *Best Pract Res Clin Gastroenterol* 19: 3-23.
5. Wu HC, Wang Q, Yang HI, Ahsan H, Tsai WY, et al. (2008) Urinary 15-F2t-isoprostane, aflatoxin B1 exposure and hepatitis B virus infection and hepatocellular carcinoma in Taiwan. *Carcinogenesis* 29: 971-976.
6. Sanyal AJ, Yoon SK, Lencioni R (2010) The etiology of hepatocellular carcinoma and consequences for treatment. *Oncologist* 15: 14-22.
7. Llovet JM, Bruix J (2000) Early diagnosis and treatment of hepatocellular carcinoma. *Baillieres Best Pract Res Clin Gastroenterol* 14: 991-1008.
8. Akriviadis EA, Llovet JM, Efremidis SC, Shouval D, Canelo R, et al. (1998) Hepatocellular carcinoma. *Br J Surg* 85: 1319-1331.
9. Hsu C, Cheng JC, Cheng AL (2004) Recent advances in non-surgical treatment for advanced hepatocellular carcinoma. *J Formos Med Assoc* 103: 483-495.
10. Zhang B, Yang B (1999) Combined alpha fetoprotein testing and ultrasonography as a screening test for primary liver cancer. *J Med Screen* 6: 108-110.
11. Masuzaki R, Karp SJ, Omata M (2012) New serum markers of hepatocellular carcinoma. *Semin Oncol* 39: 434-439.
12. Chiou SH, Wu CY (2011) Clinical proteomics: current status, challenges, and future perspectives. *Kaohsiung J Med Sci* 27: 1-14.
13. Ideker T, Thorsson V, Ranish JA, Christmas R, Buhler J, et al. (2001) Integrated genomic and proteomic analyses of a systematically perturbed metabolic network. *Science* 292: 929-934.
14. Washburn MP, Wolters D, Yates JR 3rd (2001) Large-scale analysis of the yeast proteome by multidimensional protein identification technology. *Nat Biotechnol* 19: 242-247.
15. McCormack AL, Schieltz DM, Goode B, Yang S, Barnes G, et al. (1997) Direct analysis and identification of proteins in mixtures by LC/MS/MS and database searching at the low-femtomole level. *Anal Chem* 69: 767-776.
16. MacCoss MJ, McDonald WH, Saraf A, Sadygov R, Clark JM, et al. (2002) Shotgun identification of protein modifications from protein complexes and lens tissue. *Proc Natl Acad Sci U S A* 99: 7900-7905.
17. Wu CJ, Chen YW, Tai JH, Chen SH (2011) Quantitative phosphoproteomics studies using stable isotope dimethyl labeling coupled with IMAC-HILIC-nanoLC-MS/MS for estrogen-induced transcriptional regulation. *J Proteome Res* 10: 1088-1097.
18. Santos HM, Kouvonen P, Capelo JL, Corthals GL (2012) Isotopic labelling of peptides in tissues enhances mass spectrometric profiling. *Rapid Commun Mass Spectrom* 26: 254-262.
19. Boersema PJ, Aye TT, van Veen TA, Heck AJ, Mohammed S (2008) Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. *Proteomics* 8: 4624-4632.
20. Yoshida T (1997) Peptide separation in normal phase liquid chromatography. *Anal Chem* 69: 3038-3043.
21. Yoshida T (1998) Calculation of peptide retention coefficients in normal-phase liquid chromatography. *J Chromatogr A* 808: 105-112.
22. Alpert AJ (1990) Hydrophilic-interaction chromatography for the separation of peptides, nucleic acids and other polar compounds. *J Chromatogr* 499: 177-196.
23. Perkins DN, Pappin DJ, Creasy DM, Cottrell JS (1999) Probability-based protein identification by searching sequence databases using mass spectrometry data. *Electrophoresis* 20: 3551-3567.
24. Hirosawa M, Hoshida M, Ishikawa M, Toya T (1993) MASCOT: multiple alignment system for protein sequences based on three-way dynamic programming. *Comput Appl Biosci* 9: 161-167.
25. Mi H, Muruganujan A, Thomas PD (2013) PANTHER in 2013: modeling the evolution of gene function, and other gene attributes, in the context of phylogenetic trees. *Nucleic Acids Res* 41: D377-386.
26. Nikolsky Y, Bryant J (2009) Protein networks and pathway analysis. Preface. *Methods Mol Biol* 563: v-vii.
27. Thomas PD, Kejariwal A, Guo N, Mi H, Campbell MJ, et al. (2006) Applications for protein sequence-function evolution data: mRNA/protein expression analysis and coding SNP scoring tools. *Nucleic Acids Res* 34: W645-650.
28. Unwin RD, Craven RA, Harnden P, Hanrahan S, Totty N, et al. (2003) Proteomic changes in renal cancer and co-ordinate demonstration of both the glycolytic and mitochondrial aspects of the Warburg effect. *Proteomics* 3: 1620-1632.
29. Craven RA, Stanley AJ, Hanrahan S, Dods J, Unwin R, et al. (2006) Proteomic analysis of primary cell lines identifies protein changes present in renal cell carcinoma. *Proteomics* 6: 2853-2864.
30. Wu CJ, Hsu JL, Huang SY, Chen SH (2010) Mapping N-terminus phosphorylation sites and quantitation by stable isotope dimethyl labeling. *J Am Soc Mass Spectrom* 21: 460-471.
31. Huang SY, Tsai ML, Wu CJ, Hsu JL, Ho SH, et al. (2006) Quantitation of protein phosphorylation in pregnant rat uteri using stable isotope dimethyl labeling coupled with IMAC. *Proteomics* 6: 1722-1734.
32. Thingholm TE, Jensen ON, Robinson PJ, Larsen MR (2008) SIMAC (sequential elution from IMAC), a phosphoproteomics strategy for the rapid separation of monophosphorylated from multiply phosphorylated peptides. *Mol Cell Proteomics* 7: 661-671.
33. Hou W, Ethier M, Smith JC, Sheng Y, Figeys D (2007) Multiplexed proteomic reactor for the processing of proteomic samples. *Anal Chem* 79: 39-44.
34. Pichler P, Köcher T, Holzmann J, Mazanek M, Taus T, et al. (2010) Peptide labeling with isobaric tags yields higher identification rates using iTRAQ 4-plex compared to TMT 6-plex and iTRAQ 8-plex on LTQ Orbitrap. *Anal Chem* 82: 6549-6558.
35. Webster J, Oxley D (2012) Protein identification by MALDI-TOF mass spectrometry. *Methods Mol Biol* 800: 227-240.
36. Mi H, Thomas P (2009) PANTHER pathway: an ontology-based pathway database coupled with data analysis tools. *Methods Mol Biol* 563: 123-140.
37. Chiou SH, Lee MY, Kuo CJ, Liang SS (2013) Application of quantitative proteomics to disease diagnosis and prognosis using stable isotope dimethyl labeling coupled with shotgun proteomics strategy and multiple reaction monitoring. In: Li MD (ed) *Recent Advances in Proteomics Research* (pp. 119-137), Nova Science Publishers, New York.
38. Liang SS, Kuo CJ, Chi SW, Wu WJ, Chen ST, et al. (2013) Quantitative Proteomics Analysis of Differentially Expressed Proteins Involved in Renal Clear Cell Carcinoma by Shotgun Approach Coupled with Stable Isotope Dimethyl Labeling. *J Proteomics Bioinform* 7: 003.