Reveal the molecular signatures of hepatocellular carcinoma with different sizes by iTRAQ based quantitative proteomics

Yingchao Wang a,b, Hongzhi Liu a,b,c, Dong Liang a,b,c, Yao Huang a,b,c, Yongyi Zeng a,b,c, Xiaohua Xing a,b, Jiangbao Xia a,b, Minjie Lin a,b, Xiao Han d, Naishun Liao a,b, Xiaolong Liu a,b,c, Jingfeng Liu a,b,c,⁎

a The Liver Center of Fujian Province, Fujian Medical University, Fuzhou 350025, People's Republic of China
b The United Innovation of Mengchao Hepatobiliary Technology Key Laboratory of Fujian Province, Mengchao Hepatobiliary Hospital of Fujian Medical University, Fuzhou 350025, People's Republic of China
c Liver Disease Center, The First Affiliated Hospital of Fujian Medical University, Fuzhou 350007, People's Republic of China
d Biotechnology Research Institute, Chinese Academy of Agricultural Sciences, Beijing 100081, People's Republic of China

A R T I C L E I N F O

Article history:
Received 30 March 2016
Received in revised form 3 August 2016
Accepted 19 September 2016
Available online 28 September 2016

Keywords:
Hepatocellular carcinoma
Quantitative proteomics
Tumor size
Tumor growth
iTRAQ

A B S T R A C T

Tumor size of hepatocellular carcinoma (HCC) is a key parameter for predicting prognosis of HCC patients. The biological behaviors of HCC, such as tumor growth, recurrence and metastasis are significantly associated with tumor size. However, the underlying molecular mechanisms remain unclear. Here, we applied iTRAQ-based proteomic strategy to analyze the proteome differences among small, media, large and huge primary HCC tissues. In brief, 88 proteins in small HCC, 69 proteins in medium HCC, 118 proteins in large HCC and 215 proteins in huge HCC, were identified by comparing the proteome of cancerous tissues with its corresponding non-cancerous tissues. Further analysis of dysregulated proteins involved in signaling revealed that alteration of ERK1/2 and AKT signaling played important roles in the tumorigenesis or tumor growth in all subtypes. Interestingly, alteration of specific signaling was discovered in small and huge HCC, which might reflect specific molecular mechanisms of tumor growth. Furthermore, the dysregulation degree of a group of proteins has been confirmed to be significantly correlated with the tumor size; these proteins might be potential targets for studying tumor growth of HCC. Overall, we have revealed the molecular signatures of HCC with different tumor sizes, and provided fundamental information for further in-depth study.

Biological significance: In this study, we compared the protein expression profiles among different HCC subtypes, including small HCC, media HCC, large HCC and huge HCC for the first time. The results clearly proved that different molecular alterations and specific signaling pathways were indeed involved in different HCC subtypes, which might explain the different malignancy biological behaviors. In addition, the dysregulation degree of a group of proteins has been confirmed to be significantly correlated with the tumor size. We believe that these findings would help us better understand the underlying molecular mechanisms of the tumorigenesis and development of HCC.

© 2016 Elsevier B.V. All rights reserved.

1. Introduction

Hepatocellular carcinoma (HCC) is the fifth most common malignant cancer and the second most common cause of cancer mortality worldwide [1]. Approximately 400,000 people die from HCC every year in Asian countries, which accounts for 51% of the liver cancer related deaths worldwide [2]. Although surgical resection strategies, which have been proved to be the first choice for HCC treatment, are significantly improved during the past decades, the prognosis of HCC patients is still unsatisfactory [3]. In clinical practice, it has been well established that several tumorous factors, including tumor size, portal vein tumor thrombus (PVTT), tumor lesion number, and differentiation grade of tumor, are important predictors of the prognosis of HCC patients. Among these factors, the tumor size is a crucial factor for staging the HCC patients by either Barcelona Clinic Liver Cancer (BCLC) staging classification system [4] or TNM-based staging system [5], and it is also one of the most important parameters to affect the treatment outcomes of HCC patients, such as successful rate of surgical resection, operative complications, survival, recurrence and metastasis [6–10]. According to the tumor size of patients, HCC could be divided into 4 sub-types clinically, including small HCC (diameter ≤ 3 cm), media HCC (3 cm < diameter ≤ 5 cm), large HCC (5 cm < diameter < 10 cm) and huge HCC (diameter ≥ 10 cm). The biological behaviors and clinical therapeutic outcomes of these different HCC subtypes are extremely different from each other. For
example, it has been clearly proved that HCC with larger tumor size was remarkably positively associated with higher degree of capsular invasion [11], and higher incidences of microscopic vascular invasion [7]; Ishii et al. have reported that the tumor size was the only independent risk factor for lung metastasis of HCC by analyzing the prognosis of 293 HCC patients who have undergone surgical resection, and significantly higher incidences of lung metastasis occurred in large or huge HCC patients even if there was no macroscopic vascular invasion [8]. Meanwhile, it has also been reported that the one-year recurrence rate of HCC after radical resection in patients with large or huge HCC is remarkably higher than that in patients with small or media HCC [11–15]. Furthermore, the overall survival rate, or five-year survival rate of patients has been clearly proved to be negatively associated with the tumor size [12,16]. However, the molecular signatures rather than the easily visible size features, as well as the underlying molecular mechanisms of the different biological behaviors of these different HCC subtypes are largely unknown, and need to be further elucidated.

The high-throughput quantitative proteomic strategies are ideal tools for systematically characterizing the overall proteome differences among different disease status, such as revealing the differences in protein expression or protein post-translational modifications that are caused by a particular disease status, which might provide fundamental information for in-depth understanding of complicated diseases such as cancer. Especially, the use of isobaric tags for relative and absolute quantitation (iTRAQ) profiling technology is an ultrasensitive and precise approach to simultaneously quantify and compare the proteome differences up to 8 group samples [17,18]. It has been widely used for investigating the molecular mechanisms of the tumorgenesis, tumor development, and recurrence/metastasis of HCC [19,20], as well as for the screening of diagnostic and prognostic biomarkers of HCC [21–23]. For instance, Huang et al. [23] have quantitatively compared the proteome alterations of huge HCC (diameter is larger than 10 cm) at different recurrence/metastasis stages, and discovered two biomarkers for distinguishing and predicting the early recurrence/metastasis of huge HCC through applying the iTRAQ based strategies; Xing et al. have systematically compared the proteome differences, and revealed the possible mechanism differences between the HCC with a single lesion and HCC with multiple lesions by iTRAQ based quantitative analysis [22].

However, the application of iTRAQ labeling in studying the molecular differences among primary HCC tissues with different tumor sizes has never been reported. Herein, this study aims to carefully investigate the molecular signatures of the different HCC subtypes, including small, media, large and huge HCC, by quantitatively comparing the overall proteome of the primary tumor tissues with its corresponding adjacent noncancerous tissues using iTRAQ coupling with two-dimensional liquid chromatography–tandem mass spectrometry (2D LC–MS/MS).

2. Materials and methods

2.1. Sample collection

Primary HCC tissues with different sizes were collected from 60 patients and were divided into 4 groups according to the tumor size, including small HCC group (SC, n = 15), media HCC group (MC, n = 15), large HCC group (LC, n = 15) and huge HCC group (HC, n = 15). The mean diameter of tumor was 2.2 ± 0.5 cm in the SC group, 3.9 ± 0.5 cm in the MC group, 6.9 ± 1.2 cm in the LC group, and 14.3 ± 3.4 cm in the HC group, respectively.

Furthermore, the overall survival rate, or five-year survival rate of patients has been clearly proved to be negatively associated with the tumor size [12,16]. However, the molecular signatures rather than the easily visible size features, as well as the underlying molecular mechanisms of the different biological behaviors of these different HCC subtypes are largely unknown, and need to be further elucidated.

The high-throughput quantitative proteomic strategies are ideal tools for systematically characterizing the overall proteome differences among different disease status, such as revealing the differences in protein expression or protein post-translational modifications that are caused by a particular disease status, which might provide fundamental information for in-depth understanding of complicated diseases such as cancer. Especially, the use of isobaric tags for relative and absolute quantitation (iTRAQ) profiling technology is an ultrasensitive and precise approach to simultaneously quantify and compare the proteome differences up to 8 group samples [17,18]. It has been widely used for investigating the molecular mechanisms of the tumorgenesis, tumor development, and recurrence/metastasis of HCC [19,20], as well as for the screening of diagnostic and prognostic biomarkers of HCC [21–23]. For instance, Huang et al. [23] have quantitatively compared the proteome alterations of huge HCC (diameter is larger than 10 cm) at different recurrence/metastasis stages, and discovered two biomarkers for distinguishing and predicting the early recurrence/metastasis of huge HCC through applying the iTRAQ based strategies; Xing et al. have systematically compared the proteome differences, and revealed the possible mechanism differences between the HCC with a single lesion and HCC with multiple lesions by iTRAQ based quantitative analysis [22].

However, the application of iTRAQ labeling in studying the molecular differences among primary HCC tissues with different tumor sizes has never been reported. Herein, this study aims to carefully investigate the molecular signatures of the different HCC subtypes, including small, media, large and huge HCC, by quantitatively comparing the overall proteome of the primary tumor tissues with its corresponding adjacent noncancerous tissues using iTRAQ coupling with two-dimensional liquid chromatography–tandem mass spectrometry (2D LC–MS/MS).

### Table 1

<table>
<thead>
<tr>
<th>Patient demographics and clinical characteristics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC (n = 15)</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td><strong>Age range</strong></td>
</tr>
<tr>
<td>34–71</td>
</tr>
<tr>
<td><strong>Diameter of tumor</strong></td>
</tr>
<tr>
<td>2.2 ± 0.5</td>
</tr>
<tr>
<td><strong>AFP</strong></td>
</tr>
<tr>
<td>&gt;10</td>
</tr>
<tr>
<td>≤10</td>
</tr>
<tr>
<td><strong>TNM stage</strong></td>
</tr>
<tr>
<td>I/II</td>
</tr>
<tr>
<td>III/IV</td>
</tr>
<tr>
<td><strong>Liver cirrhosis</strong></td>
</tr>
<tr>
<td>None</td>
</tr>
<tr>
<td>Mild/moderate</td>
</tr>
<tr>
<td>Severe</td>
</tr>
</tbody>
</table>

Note: AFP, alpha-fetoprotein. AFP value >10 is positive through the manufacturer’s introduction. Two patients in the LC group and one patient in the HC group didn’t accept the AFP examination, respectively.

3.4. Large HCC group (LC, n = 15) and huge HCC group (HC, n = 15).

### Table 1

<table>
<thead>
<tr>
<th>Patient demographics and clinical characteristics.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SC (n = 15)</td>
</tr>
<tr>
<td>---</td>
</tr>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td><strong>Age range</strong></td>
</tr>
<tr>
<td>34–71</td>
</tr>
<tr>
<td><strong>Diameter of tumor</strong></td>
</tr>
<tr>
<td>2.2 ± 0.5</td>
</tr>
<tr>
<td><strong>AFP</strong></td>
</tr>
<tr>
<td>&gt;10</td>
</tr>
<tr>
<td>≤10</td>
</tr>
<tr>
<td><strong>TNM stage</strong></td>
</tr>
<tr>
<td>I/II</td>
</tr>
<tr>
<td>III/IV</td>
</tr>
<tr>
<td><strong>Liver cirrhosis</strong></td>
</tr>
<tr>
<td>None</td>
</tr>
<tr>
<td>Mild/moderate</td>
</tr>
<tr>
<td>Severe</td>
</tr>
</tbody>
</table>

Note: AFP, alpha-fetoprotein. AFP value >10 is positive through the manufacturer’s introduction. Two patients in the LC group and one patient in the HC group didn’t accept the AFP examination, respectively.

### 2.2. Protein preparation and iTRAQ labeling

Protein preparation, peptide labeling and 2D LC–MS/MS analysis were performed following the previously published protocols [22] with slight modification. In brief, the tissues from patients were divided into 8 groups as above-mentioned, including SC paired with SN, MC paired with MN, LC paired with LN and HC paired with HN, respectively.

To extract the proteins from tissues, 300 μL of lysis buffer containing 8 M urea, 2% SDS and 1× Protease Inhibitor Cocktail (Roche Ltd. Switzerland) was added into the mixed samples, then followed by tissue homogenization and sonication on ice. After centrifugation at 17,000g for 10 min at 4 °C, the supernatant was collected and transferred to a fresh tube. The protein concentration of the raw extraction was determined by the standard radical resection [24]: no distal metastasis was revealed in both pre- and intra-operative examinations; no lesion was found in the rest of the liver during intra-operative ultrasonic scan; no visible cancer embolus in the hepatic portal vein or primary venous branch; no cancer cell was found in the incisal margin at the post-operative pathological examinations, the encapsulation of tumor tissue was intact and the boundary of tumor tissue was distinct; and no recurrent/metastatic lesion was found at the ultrasonic and CT scan during the return visit after 2 months of surgery; (4) The elevated pre-operative serum AFP should decline to normal level after 2 months post-operation; and (5) The patient did not undergo any other intervention or therapies before surgery.

Fresh tissues were collected at the time of surgery from patients with Hepatitis B virus (HBV) associated primary HCC; part of the collected tissues was immediately liquid nitrogen preserved after washing with phosphate-buffered saline (PBS), and part of the tissues was formalin embedded and stored for immunohistochemistry. The histological diagnosis of the tissue samples was confirmed by experienced pathologists. The project was approved for the using of human biopsy by the Institution Review Board of Mengzho Hepatobiliary Hospital of Fujian Medical University. The written consent was received from all participants in this study.
group to perform the technical and biological repeats to ensure data quality and reliability. The final volume of protein mixtures was adjusted to 300 μL with 100 mM TEAB (triethylammonium bicarbonate, Santa Cruz, USA). Afterwards, the protein mixtures were reduced and alkylated by 5 μL DTT (200 mM) and 10 μL iodoacetamide (500 mM) through our previous protocols [22,23], respectively.

The proteins from each mixture were precipitated by ice-cold acetone, and then were re-dissolved in 100 μL TEAB (100 mM). Afterwards, the proteins were typically digested by sequence-grade modified trypsin (Promega, Madison, WI), and then the resultant peptide mixture was further labeled using chemicals from the iTRAQ reagent kit (AB SCIEX, USA) as follows: the SC and SN were labeled with 116 and 114 isobaric tags, respectively; the MC and MN were labeled with 115 and 113 isobaric tags, respectively; the LC and LN were labeled with 117 and 121 isobaric tags, respectively; and the HC and HN were labeled with 118 and 119 isobaric tags, respectively. Each isobaric tag labeled sample was pooled together, and the iTRAQ 8-plex labeling was independently repeated 3 times, defining as Exp A, Exp B and Exp C. Finally, the labeled samples were desalted with the Sep-Pak Vac C18 cartridges (Waters Corporation, USA) and then dried in a vacuum centrifuge for further usage.

2.3. High pH reverse phase separation

The peptide mixture was re-dissolved in the buffer A (buffer A: 20 mM ammonium formate in water, pH 10.0 which was adjusted with ammonium hydroxide), and then fractionated by high pH separation using an Acquity UPLC system (Waters Corporation, Milford, MA) connected to a reverse phase column (XEbridge C18, 3.5 μm, 2.1 mm × 150 mm, Waters Corporation, USA). High pH separation was performed using a linear gradient starting from 5% buffer B to 35% buffer B in 45 min (buffer B: 20 mM ammonium formate in 90% ACN, pH 10.0 which was adjusted with ammonium hydroxide). The column flow rate was maintained at 200 μL/min and column temperature was maintained at room temperature. After the separation, the column was re-equilibrated at initial conditions for 15 min. Finally 20 fractions were collected, and each fraction was dried in a vacuum concentrator for the next step.

2.4. Low pH nano-LC-MS/MS analysis

The fractions were re-suspended with 32 μL buffer C (buffer C: 0.1% formic acid in dd-water), separated by nano-LC, and then were analyzed by on-line electrospray tandem mass spectrometry. The experiments were implemented on a Nano-Accuity UPLC system (Waters Corporation, USA) connected to a quadrupole-Orbitrap mass spectrometer (Q-Exactive) (Thermo Fisher Scientific, Germany) equipped with an online nano-electrospray ion source. 8 μL of peptide sample was loaded onto the trap column (Thermo Scientific Acclaim PepMap C18, 100 μm × 2 cm) with a flow rate of 10 μL/min for 3 min, and subsequently was separated on the analytical column (Acclaim PepMap C18, 75 μm × 50 cm) with a linear gradient, from 2% to 40% buffer D (buffer D: ACN including 0.1% formic acid) in 135 min. The column flow rate was kept at 300 nL/min and the column temperature was maintained at 40 °C. The electrospray voltage of 2.2 kV versus the inlet of the mass spectrometer was used. After the nano-LC separation, the column was re-equilibrated at initial conditions for 15 min. The Q-Exactive mass spectrometer was operated in the data-dependent mode to switch automatically between MS and MS/MS acquisition. Survey full-scan MS spectra (m/z 350–1200) were obtained with a mass resolution of 70 K, followed by fifteen sequential high energy collisional dissociation (HCD) MS/MS scans with a resolution of 17.5 K. In all cases, one microscan was recorded by dynamic exclusion of 30 s.

2.5. Data analysis

Data analysis was performed according to our previous report [22]. In brief, the raw files were exported into mzXML and MGF format files with MS convert module in Trans-Proteomic Pipeline (TPP 4.6.2). The raw data was searched by Thermo Scientific Proteome Discoverer software version 1.4 with the MASCOT (Matrix Science, London, UK; version 2.3.2) search engine against human database provided by The Universal Protein Resource (http://www.uniprot.org/uniprot, released at 2014-04-10, with 20,264 entries). The key parameters for searching were set as follows: (1) protein sequence was trypsinized specifically; (2) allow up to 2 missed cleavages maximally; (3) allow a parent ion tolerance of 10 parts per million (ppm) and a fragment ion mass tolerance of 0.05 Da; (4) allow fixed modification including carboximidemethylation of cysteine, iTRAQ modification of peptide N-terminus and lysine residues; and (5) allow variable modifications including oxidation of methionine and iTRAQ 8-plex labeling of tyrosine. Scaffold software (version Scaffold_4.3.2, Proteome Software Inc., Portland, OR) was used to supervise the Identification Quality. The identification of proteins was accepted when the false discovery rate (FDR) of peptide is <1% and protein probability is higher than 99.0%.

The differentially expressed proteins were selected by following settings: the fold-change of protein expressions in cancerous tissues (compared to its non-cancerous tissues) was higher than 2 for up-regulation or was lower than 0.5 for down-regulation, with the p < 0.05, respectively, statistical analysis of which was performed by the paired t-test.

2.6. Bioinformatics analysis and hierarchical clustering of differentially expressed proteins

The Gene Ontology (GO) annotation and classification of differentially expressed proteins were implemented by PANTHER (Protein Analysis through Evolutionary Relationships) Classification System (http://www.pantherdb.org/). Ingenuity Pathways Analysis (IPA) software (version 7.5) was used to analyze the biological functions and signaling pathways. The identified proteins were enriched to the most significant networks generated from previous publications and public protein interaction databases. A p-value calculated with the right-tailed Fisher’s exact test was used to score and to rank networks according to their degree of association with differentially expressed proteins. The differentially expressed proteins that significantly correlated with the tumor size were hierarchically clustered by R software (version: i386 3.1.1) by loading “gplots” program. The differentially expressed proteins were subjected to hierarchical cluster analysis with un-centered correlation used as similarity metric and complete linkage as clustering method.

2.7. qRT-PCR analysis of differentially expressed proteins

To further confirm the expression level alterations of identified proteins, the quantitative RT-PCR (qRT-PCR) assays were performed in tumor tissues and its corresponding noncancerous, as previously described with slight modification [25]. Briefly, total RNA was extracted from fresh-frozen tissues using Trizol reagent (Invitrogen, CA) following the manufacturer’s recommendation, and quantified using Nanodrop spectrophotometry (Thermo Scientific, USA). Reverse transcription was performed by using the GoScript Reverse Transcription System kit (Promega, USA). The qRT-PCR was run in technical duplicates for each reaction with 50 ng cDNA from triplicate of each sample through using GoTag qPCR Master mix (Promega, USA). Relevant information on gene-specific primers used to detect the gene expression level of identified proteins was designed using Primer-Blast Database, and was shown in Supplementary Table S1. All primers used in this study were verified with PCR amplification efficiency which was ranked between 90% and 110%. The cycling parameters were 40 cycles of 95 °C for 15 s, 60 °C for 60 s. The GeNorm applet (v 3.5) was used to select the most
2.8. Immunoblotting analysis of differentially expressed proteins

To further confirm the protein expression level change of selected proteins, western blot (WB) was performed in 9 pairs of cancerous and non-cancerous tissues in each subgroup for verification. The tissues from individual patients were homogenated and then lysed by RIPA buffer on ice for 30 min [26]. Afterwards, the protein extraction was quantified by BCA assay according to the manufacturer’s instruction. Then, 50 μg protein of each sample was separated by SDS-PAGE and transferred onto the NC membranes (PALL Corporation, USA). Subsequently, the membranes containing the separated proteins were blocked with 5% BSA in TBST buffer (25 mM Tris, 190 mM NaCl, 0.05% Tween 20, pH 7.5) for 2 h at room temperature, and probed with the FGG (1:1000, Santa Cruz Biotech, USA), HK2 (1:1000, Abcam, USA), and β-actin (1:2000, Santa Cruz Biotech, USA) primary antibody at 4 °C overnight. After washing 3 times with TBST buffer for 10 min of each, the membranes were incubated with appropriate horseradish peroxidase (HRP) conjugated secondary antibodies (1:5000 dilution, TransGen Biotech, China) for 1 h at room temperature. Following 3 times wash again in TBST buffer, the blots were detected by enhanced chemiluminescence and visualized by the ChemiDoc MP imaging system (Bio-Rad, Hercules, CA). Quantification of target proteins was analyzed by densitometry and normalized by β-actin.

2.9. Statistical analysis

Prism statistical software (GraphPad v6.01, CA) and SPSS statistics (v 19.0, SPSS Inc., USA) were employed for data analysis. Quantitative variables are presented as means ± SD. Correlation analysis was performed using two-tailed Pearson coefficient. p < 0.05 was considered as statistically significant.

3. Results

3.1. Proteomic analysis of HCC subtypes with different tumor sizes

An experimental strategy relying on the use of quantitative proteomics approach (iTRAQ) was designed to compare the overall proteome differences among primary HCC with different sizes. In clinical practice, the HCC could be divided into 4 subtypes according to tumor diameter. In this study, the clinical samples were categorized into 8 sub-groups as mentioned in the “Sample collection” section, including SC paired with SN; MC paired with MN; LC paired with LN; and HC paired with HN. All of the patients were selected with similar age distributions, sex distributions, differentiation degrees, degree of tumor boundaries definition, degree of tumor encapsulation integrity, and similar vascular tumor thrombosis. The workflow of the whole study is shown in Fig. 1. The mix strategy of samples as mentioned in the “Materials and methods” section could allow to narrow down the individual differences among patients.

3.2. Relative quantification of proteomes of different HCC subtypes

Using the analysis results from software Scaffold, 4842, 4863 and 4801 proteins were identified in three iTRAQ 8-plex labeling replicates respectively, with the following criteria: at least 1 tryptic peptide of >95% confidence, and false discovery rate (FDR) <1%. And 4277 proteins were shared by three replicates, accounting for 81.9% the total quantified proteins (Fig. 2A). In the present study, we have excluded the single or double-hit identified proteins, and only took the proteins which were presented in all 3 replicas for further analysis. The complete list of identified proteins and their characteristics including molecular weight (MW), isoelectric point (pI), hydrophobicity, exponentially modified Protein Abundance Index (emPAI), average coefficient of variance (CV), protein grouping ambiguity, number of identified peptides, PSMs, quantitative clustering, and quantification results with percentage variability, are available in Table S2. The raw proteomics data have been deposited into integrated Proteome resources and can be accessed through http://www.iprox.org/index (the dataset is available for private download at that website via username: reviewer509 and password: zpv6pvrf).

The MW, PI and hydrophobicity of identified proteins were well distributed at common ranges, indicating that the overall proteome datasets of the different HCC subtypes had no strong bias (Fig. S1). The abundance level of all the quantified proteins [27] in triplicates were compared by calculating their emPAI (Fig. 2B). The abundance was mainly distributed in range from 2 to −3, and exhibited well repeatability. The subcellular localizations of the identified proteins were also analyzed by Gene Ontology annotation (Fig. S1). Although cytoplasmic protein (25.9%) and membranous proteins (15.99%) were the most highly represented categories in our extracts, the nucleoprotein (12.48%) and protein complex (9.39%) were also readily identified; it suggests that our protein extraction procedure was not strongly biased to a few cell compartments.

3.3. Differentially expressed proteins between cancerous and non-cancerous tissues of different HCC subtypes

In total, 88 differentially expressed proteins (34 up- and 54 down-regulated) in the SC group (Fig. 2C), 69 differentially expressed proteins
(44 up- and 25 down-regulated) in the MC group (Fig. 2D), 118 differentially expressed proteins (66 up- and 52 down-regulated) in the LC group (Fig. 2E), and 215 differentially expressed proteins (139 up- and 76 down-regulated) in the HC group (Fig. 2F), were identified by mean expression fold change of ±2 folds with p < 0.05 (paired t-test) by comparing the protein expression between cancerous tissues and non-cancerous tissues. Meanwhile these proteins should have the same change trends in all 3 biological replicates. The full information of these differentially expressed proteins of different HCC subtypes were listed in Table S3.

In GO classification analysis, the dysregulated proteins involved in biological processes (Fig. 2G) in all subtypes are mainly focused on metabolic process, cellular process, protein localization, development, immune responses as well as apoptotic process. In the clustering of molecular functions (Fig. 2H), the majority of aberrantly expressed proteins were involved in nucleic acid binding, oxidoreductase, transferase, hydrolase, ligase, enzyme activity modulation and protease functions. These results revealed that extensive changes of biological processes and molecular functions have occurred inside the tumors, and the change levels were gradually increased along with the tumor size. Therefore, these changes might be important for the tumorigenesis and development of HCC. Among these differentially expressed proteins, 6 proteins including MT1G, FHDC1, TTC36, PBLD, AKR1B10, HIST1H1C, and HIST1H1B, significantly altered their expressions in the tumor tissues of all HCC subtypes.

3.4. IPA network analysis of the differentially expressed proteins in each HCC subtype

The involved signaling pathways or interaction networks of differentially expressed proteins in each HCC subtype were analyzed by IPA.
The IPA analysis showed that the dysregulated proteins in cancerous tissues of patients with small HCC were mostly involved in ERK1/2 (Fig. 3A), AKT (Fig. S2A), MAPK (Fig. S2A), and UBC (Fig. S2B) signaling pathways; the dysregulated proteins in patients with media HCC were mainly involved in ERK1/2 (Fig. 3B), MAPK (Fig. S2C) and AKT (Fig. S2D) signaling pathways; the dysregulated proteins in patients with large HCC were involved in ERK1/2 (Fig. 3C), AKT (Fig. S2E), MAPK (Fig. S2F) signaling pathways; while the dysregulated proteins in patients with huge HCC were mainly concentrated in NF-κB (Fig. 3D), CDK2 (Fig. S2G), AKT (Fig. S2H) and ERK1/2 (data not shown) signaling pathways.

According to the IPA analysis, the ERK1/2 and AKT signaling pathways are generally involved in all HCC subtypes; especially, the ERK1/2 signaling pathway obtained the top score and enriched 25, 19 and 26 dysregulated proteins in the SC group, MC group and LC group, respectively; in the HC group, 18 dysregulated proteins were enriched in ERK1/2 pathway, although which did not obtain the highest score. Meanwhile, AKT signaling pathway assembled 18, 17, 24 and 19 dysregulated proteins in SC, MC, LC and HC groups, respectively. It has been well confirmed that the ERK1/2 and AKT signaling pathways played extremely important roles in cancer cell growth, proliferation and anti-apoptosis [28–31]. Therefore, it is not surprising that the ERK1/2 and AKT signaling pathways were involved in all HCC subtypes.

Interestingly, only UBC signaling pathway (15 dysregulated proteins were enrolled) was specifically involved in SC group. This result was supported by a published report, which has shown the abnormality of UBC system specifically in HCC ≤ 3 cm (small HCC) [32]. It has been reported that the UBC pathway played an irreplaceable role in the ubiquitin–proteasome proteolytic system in mammals, and precisely controlled 80–90% of protein degradation [33]. Our discovery indicated that a serious disorder of protein degradation has occurred in small HCC. Although aberrations of UBC pathway have been implicated in the pathogenesis of several cancers [33–35], the underlying mechanisms of the disturbance of UBC signaling pathway specifically in small HCC, remain unclear and need further investigation.

Similarly, NF-κB (31 dysregulated proteins were accumulated) and CDK2 (25 dysregulated proteins were enrolled) signaling pathways were only specifically emerged in the HC group. It is well known that NF-κB is an essential transcriptional factor in charge of enhancing tumor angiogenesis, proliferation, anti-apoptosis, and repression of immune response [36,37]; meanwhile, CDK2 signaling plays a crucial role in cancer cell growth, proliferation, and cell cycle progression.

![Fig. 3.](image-url)
in regulating various events of eukaryotic cell cycles, which was directly associated with the proliferation of cancer cells [38,39]. The aberrations of NF-κB and CDK2 signaling pathways are in good agreement with the stronger growth ability of huge HCC.

Overall, the above analysis revealed that the alteration of specific signaling pathways is indeed involved in the tumorigenesis and development of small HCC and huge HCC, although some common signaling pathways emerged in all HCC subtypes. Meanwhile, it also clearly proves that the proteomics based approach might be a suitable tool for investigating the overall molecular profile changes, as well as the underlying molecular mechanisms of the different HCC subtypes.

3.5. Hierarchal clustering of the differentially expressed proteins in tumor of different HCC subtypes

To further understand the molecular signatures associated with the tumor size or tumor growth of HCC, the hierarchal molecular clustering of 57 dysregulated proteins in the tumor of different HCC subtypes was performed. As shown in Fig. 4, two types of distinct clusters were clearly identified; the degree of upregulation of 44 proteins (cluster 1) was significantly positively associated with bigger tumor size; while the degree of downregulation of 13 proteins (cluster 2) was significantly negatively associated with bigger tumor size. The names of these dysregulated proteins were listed in Table S4. Among these 57 dysregulated proteins, 23 proteins (Table 2) have been reported to be associated with the proliferation of cancer cells. Since the active degree of cell proliferation has been well established to correlate with tumor size or tumor growth of HCC, the hierarchal molecular clustering of 57 dysregulated proteins in the tumor of different HCC subtypes was performed. As shown in Fig. 4, two types of distinct clusters were clearly identified; the degree of upregulation of 44 proteins (cluster 1) was significantly positively associated with bigger tumor size; while the degree of downregulation of 13 proteins (cluster 2) was significantly negatively associated with bigger tumor size. The names of these dysregulated proteins were listed in Table S4. Among these 57 dysregulated proteins, 23 proteins (Table 2) have been reported to be associated with the proliferation of cancer cells. Since the active degree of cell proliferation has been well established to correlate with tumor size or tumor growth of HCC, the hierarchal molecular clustering of 57 dysregulated proteins in the tumor of different HCC subtypes was performed. As shown in Fig. 4, two types of distinct clusters were clearly identified; the degree of upregulation of 44 proteins (cluster 1) was significantly positively associated with bigger tumor size; while the degree of downregulation of 13 proteins (cluster 2) was significantly negatively associated with bigger tumor size. These results are well consistent with the MS (mass spectrometry) results and as shown in Fig. S3. Interestingly, these 8 proteins have been

Meanwhile, the upregulation or downregulation degrees of both cluster 1 and cluster 2 are quite similar in the SC and MC groups, but very different with the LC and HC groups. Especially, the pattern of both cluster 1 and cluster 2 is extremely different in the HC group compared with the other 3 groups. These results might reflect the unique molecular signatures and mechanisms in huge HCC of its distinct biological behaviors.

3.6. Validation of differential expression of the dysregulated proteins

According to the molecular clustering analysis, the changing degree of expression of 8 proteins was observed to be significantly correlated with the tumor size, including FGG, HK2, MT1G, PDK1, CLIC1, CKB, DMBT1 and MCM3 (Fig. S3). To further validate the observed correlations, the mRNA expression of these 8 proteins was analyzed by qRT-PCR in a large scale of samples from different subgroups including SC (n = 21), MC (n = 21), LC (n = 21), and HC (n = 21), and their corresponding non-cancerous tissues (21 samples of each subgroup), with totally 84 cancerous tissues and 84 non-cancerous tissues. Moreover, each individual cancerous tissue was detected independently and was compared versus its corresponding non-cancerous tissue without pooling. As shown in Fig. 5, the changing degrees of FGG (Fig. 5A), HK2 (Fig. 5B), PDK1 (Fig. 5D), CLIC1 (Fig. 5E), CKB (Fig. 5F), DMBT1 (Fig. 5G) and MCM3 (Fig. 5H) at mRNA level were positively correlated with tumor size (higher degree of upregulation is associated with larger tumor size); while the changing degree of MT1G mRNA (Fig. 5C) was negatively correlated with the tumor size (higher degree of downregulation is associated with larger tumor size). These results are well consistent with the MS (mass spectrometry) results as shown in Fig. S3. Interestingly, these 8 proteins have been
HCC patients such as operative complications, recurrence/metastasis, types, including small HCC, medium HCC, large HCC and huge HCC for compared the overall proteome differences among different HCC sub-

As reported in previous studies, the tumor size as an independent prognostic factor always significantly affects therapeutic outcomes of HCC patients such as operative complications, recurrence/metastasis, and overall survival. In this study, we systematically quantified and compared the overall proteome differences among different HCC subtypes, including small HCC, media HCC, large HCC and huge HCC for the first time. In total, 5217 proteins were identified with FDR < 1%, and 81.9% of these proteins were shared in all 3 biological and technical repeats. Comparing other proteomic studies in liver cancer, the protein numbers identified in our experiments are relatively high, and the high overlapping percentage of 3 experiments proved the stability of the workflow and the reliability of the research conclusion [23,50].

There are significant differences of the molecular behaviors among different HCC subtypes through comprehensive analysis of the differentially expressed proteins involved in signaling pathways. In small HCC, the generation of pro-inflammatory cytokines, chemokines, and even immunoglobulins was remarkably modulated to struggle for tumor survival in the liver (Fig. S2A). In media HCC, dysregulation of caspase 3 and TNF which play key roles in cell apoptosis [51-53], was concentrated as two central nodes; it is suggesting the anti-apoptosis events (Fig. S2C) of cancer cells were starting to play crucial roles along with the growth of tumor size. Large HCC tended to regulate the tumor growth by modulating the secretion of growth hormones, since 7 dysregulated proteins were specifically involved in this signaling pathway (Fig. 3C). The specific modulation of CDK2 signaling pathway was discovered in huge HCC; this might suggest that the change of cell cycle has happened in huge HCC (Fig. S2G). These results clearly revealed that each HCC subtype has its own molecular and biological characteristics.

In this study, we successfully filtered out 57 differentially expressed proteins, whose expression degrees change between tumor tissues and non-tumor tissues were significantly correlated with tumor sizes. These proteins could form 2 distinct clusters, which might be further applied as the molecular signatures of different HCC subtypes for molecular classification or in-depth molecular mechanism study. Among these proteins, 8 proteins which are reported to be associated with tumor growth in other cancers have been further analyzed; the higher upregulation degrees of FGG, HK2, PDK1, CKB, CLIC1, MCM3, and DMBT1, which have been reported as carcinogenic factors, were discovered in larger size tumors; while the stronger downregulation degree of MT1G, which has been reported as tumor suppressor, was discovered in smaller size tumors. These proteins might promote HCC growth in a synergistic manner. For example, MT1G, PDK1, CKB and HK2 have been reported to exert their functions probably through modulating AKT signaling pathway; DMBT1 and CLIC1 have been reported to regulate the NF-κB signaling pathway; FGG and MCM3 have been reported to be involved in the ERK 1/2 signaling pathway. All of these associated pathways have been well known to actively regulate normal cellular functions such as cell growth, cell apoptosis, as well as different aspects of cancer ranging from tumorigenesis to tumor metastasis. It is not surprising that these proteins are tightly correlated with tumor size to

4. Discussion

As reported in previous studies, the tumor size as an independent prognostic factor always significantly affects therapeutic outcomes of HCC patients such as operative complications, recurrence/metastasis, and overall survival. In this study, we systematically quantified and compared the overall proteome differences among different HCC subtypes, including small HCC, media HCC, large HCC and huge HCC for the first time. In total, 5217 proteins were identified with FDR < 1%, and 81.9% of these proteins were shared in all 3 biological and technical repeats. Comparing other proteomic studies in liver cancer, the protein numbers identified in our experiments are relatively high, and the high overlapping percentage of 3 experiments proved the stability of the workflow and the reliability of the research conclusion [23,50].
reflect the molecular basis of the different biological behaviors of different HCC subtypes. Interestingly, among all of the verified proteins, the expression levels of 3 proteins have also been identified to be correlated with the tumor size (Fig. S4, with r value \( \leq 0.9 \)), but without statistical significance, including XPO5 [54], BYSL [55,56], ELANE [57], which have been also widely studied in bladder cancer, gastric cancer, lung cancer and HCC, and have been reported to be closely associated with the proliferation and apoptosis of cancer. Therefore, it would be extremely helpful to further investigate these protein expression patterns in a large sample size. Taken together, these proteins might serve as molecular signatures of different HCC subtypes, but in-depth studies are required for better understanding the underlying molecular mechanisms.

Noticeably, the expression levels of some proteins are varied significantly in different samples even inside the same group with appearance of both upregulation and downregulation. This phenomena might be due to the following 2 reasons: (1) it might reflect the individual differences of each patient; there are also plenty of other factors that could affect the protein expression level, such as patient’s genotype, lifestyle, and subtype of HBVs (Table S5); (2) different individual patients might have different inflammatory and liver cirrhosis background to develop the HCC, which would significantly affect the protein expression ration between the para-carcinoma tissues and carcinoma tissues. Although most of these irregular dysregulation of proteins majorly happened in small or media HCC subgroups, the expressions of these proteins trended to be uniform in large HCC and huge HCC subgroups. These results might suggest that the dysregulation degree of selected proteins is significantly correlated with the degree of disease progress and severity, which is consistent with our hypothesis.
Hexokinase-2 is a glycolytic enzyme and is responsible for initiating anaerobic glucose metabolism. It has been extensively reported that HK2 was overexpressed in various cancers, such as glioblastoma [61], HCC [62], and gastric carcinoma [63]. The hypoxia microenvironment inside the tumor requires cancer cells to grow preferentially by anaerobic manner. With the increasing the tumor size, the hypoxia microenvironment inside the tumor, especially in the center of tumor, would be more severe. To survive in such a nasty environment, the cancer cells have to change from aerobic metabolism to anaerobic metabolism to adapt to the environment. Therefore, this might well explain the upregulation of HK2 in tumors, as well as explain our findings that the degree of HK2 upregulation was positively correlated with the tumor size (Pearson $r = 0.9972$ for FGG, Pearson $r = 0.9542$, $p < 0.05$), and these results are well consistent with the mRNA and MS results. NS means non-significant.

The adjacent noncancerous tissues compose the microenvironment for tumor survival and development, which is well known as the “soil” and “seeds”. Here, we also comprehensively analyzed the molecular differences among noncancerous tissues of different HCC subtypes, and were trying to identify molecular signatures and specific molecular events or mechanisms of the “soil”, that might facilitate the tumor growth or development. Surprisingly, the protein expression patterns among noncancerous tissues of different HCC subtypes were very

---

Fig. 6. Validation of the differential expression of identified proteins by WB. (A, B) The western blotting results of the FGG (A) and HK2 (B) expression in tumor tissues and their corresponding non-tumor tissues in different HCC subtypes. (C, D) The fold change of FGG (C) and HK2 (D) expression at protein level. The expression changes of FGG and HK2 are nicely correlated with the tumor size (Pearson $r = 0.9972$ for FGG, Pearson $r = 0.9542$, $p < 0.05$), and these results are well consistent with the mRNA and MS results. NS means non-significant.
similar (data not shown), and little significant differences have been identified in our experimental sets following the same criteria that previously applied. These results indicate that the background or microenvironment of different HCC subtypes is very similar, and the different tumor sizes might be predominantly due to the “seed” itself at least in our experimental settings, but it does need further confirmation.

Overall, we have applied the iTRAQ based quantitative proteomics approach to systematically compare the proteome differences among HCC subtypes with different tumor sizes, and identified possible molecular signatures of different HCC subtypes, that might be worth for further in-depth investigation.

5. Conclusions

Here, the iTRAQ-based quantitative proteomics approach has been applied to comprehensively analyze the molecular characteristics of different HCC subtypes, including small HCC, media HCC, large HCC, and huge HCC. Our results clearly proved that different molecular alterations and specific signaling pathways, which might explain the different malignancy biological behaviors, were indeed involved in different HCC subtypes. Meanwhile, a spectrum of proteins has been identified to be tightly associated with tumor size; 8 of these proteins were further confirmed at mRNA level, and 2 of these proteins were further confirmed at both mRNA and protein levels. Therefore, these proteins might serve as the molecular signatures of different HCC subtypes, and also might be interesting targets for investigating and preventing the HCC growth.

Abbreviations

HCC hepatocellular carcinoma
2DC–MS/MS two-dimensional liquid chromatography–tandem mass spectrometry
TEAB triethylammonium bicarbonate
IAA iodoacetamide
FDR false discovery rate
GO Gene Ontology
IPA ingenuity pathway analysis
MW molecular weight
emPAI exponentially modified Protein Abundance Index
RIPA lysis buffer Radio Immunoprecipitation Assay lysis buffer
TNM tumor node metastasis
CDK2 Cyclin-dependent kinase 2
MT1G Metallothionein 1G
FGG fibrinogen gamma chain
HK2 hexokinase 2
PDK1 pyruvate dehydrogenase kinase isozyme 1
MCM3 DNA replication licensing factor MCM3
CLIC1 chloride intracellular channel protein 1
DMBT1 deleted in malignant brain tumors 1
CKB creatine kinase B
ATXN10 Ataxin-10
FUS RNA-binding protein FUS
TMM50 mitochondrial inner membrane translocase subunit TMM50
TPM1 tropomyosin alpha-1 chain

Supplementary data to this article can be found online at http://dx.doi.org/10.1016/j.jprot.2016.09.008.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Grant No. 81602102); the specialized Science and Technology Key Project of Fujian Province (Grant No. 2013Y20002–3); the Science and Technology Infrastructure Construction Program of Fujian Province (Grant No. 2014Y2005); the Natural Science Foundation of Fujian Province (Grant No. 2015J05174); the Scientific research project of health and Family Planning Commission of Fujian province (Grant No. 2015–1-94); and the Scientific Foundation of Fuzhou Health Department (Grant No. 2014–S-139–3, Grant No. 2014–S-wp1, and Grant No. 2013–S-wp1).

References


Transparency document

The Transparency document associated with this article can be found in online version.
M. Kawada, H. Seno, K. Kanda, Y. Nakanishi, R. Akitake, H. Komekado, K. Kawada, Y.
N. Zheng, X. Ding, A. Sun, R. Jahan, PDK1 activity regulates proliferation, invasion
Y. Tian, Y. Guan, Y. Jia, Q. Meng, J. Yang, Chloride intracellular channel 1 regulates
J. Chen, S. Zhang, Y. Li, Z. Tang, W. Kong, Hexokinase 2 overexpression promotes the
Y. Ishihama, Y. Oda, T. Tabata, T. Sato, T. Nagasu, J. Rappsilber, M. Mann, Exponential-
S. Holck, H.J. Nielsen, N. Pedersen, L.I. Larsson, Phospho-ERK1/2 levels in cancer cell
Y. Chen, H. Wu, X.Z. Shen, The ubiquitin-proteome system and its potential application
X. Deschenses-Simard, F. Fedtowsky, S. Meloche, G. Ferbeyre, ERKs in cancer: friends or
M. Peach, N. Marsh, E.I. Miskiwicz, D.J. MacPhie, Solubilization of proteins: the impor-
Y. Iishimasa, O. Tabata, T. Sato, T. Nagaus, J. Rappislder, M. Mann, Exponential-
J. Chen, S. Zhang, Y. Li, Z. Tang, W. Kong, Hexokinase 2 overexpression promotes the
Y. Ishihama, Y. Oda, T. Tabata, T. Sato, T. Nagasu, J. Rappsilber, M. Mann, Exponential-
Y. Tian, Y. Guan, Y. Jia, Q. Meng, J. Yang, Chloride intracellular channel 1 regulates