Identification of Protein Abundance Changes in Hepatocellular Carcinoma Tissues Using PCT–SWATH

Yi Zhu, Jiang Zhu, Cong Lu, Qiushi Zhang, Wei Xie, Ping Sun, Xiaochuan Dong, Liang Yue, Yaoting Sun, Xiao Yi, Tiansheng Zhu, Guan Ruan, Ruedi Aebersold,* Shi’ang Huang,* and Tiannan Guo*

Purpose: To rapidly identify protein abundance changes in biopsy-level fresh-frozen hepatocellular carcinoma (HCC).

Experimental design: The pressure-cycling technology (PCT) is applied and sequential window acquisition of all theoretical mass spectra (SWATH-MS) workflow is optimized to analyze 38 biopsy-level tissue samples from 19 HCC patients. Each proteome is analyzed with 45 min LC gradient. MCM7 is validated using immunohistochemistry (IHC).

Results: A total of 11 787 proteotypic peptides from 2579 SwissProt proteins are quantified with high confidence. The coefficient of variation (CV) of peptide yield using PCT is 32.9%, and the $R^2$ of peptide quantification is 0.9729. Five hundred forty-one proteins showed significant abundance change between the tumor area and its adjacent benign area. From 24 upregulated pathways and 13 suppressed ones, enhanced biomolecule synthesis and suppressed small molecular metabolism in liver tumor tissues are observed. Protein changes based on $\alpha$-fetoprotein expression and hepatitis B virus infection are further analyzed. The data altogether highlight 16 promising tumor marker candidates. The upregulation of minichromosome maintenance complex component 7 (MCM7) is further observed in multiple HCC tumor tissues by IHC.

Conclusions and clinical relevance: The practicality of rapid proteomic analysis of biopsy-level fresh-frozen HCC tissue samples with PCT–SWATH has been demonstrated and promising tumor marker candidates including MCM7 are identified.

1. Introduction

Hepatocellular carcinoma is the fifth frequent malignancy worldwide and ranks as the third leading cause of cancer-related mortality.[1] In China, HCC leads to more than 300 000 deaths every year.[2] In early-stage disease, HCC patients’ survival can be significantly improved by treatment including surgical resection and liver transplant. However, HCC mostly exhibits no symptom...
at early stages. Consequently, less than one-third of newly diagnosed HCC patients are eligible for potential curative therapies. Therefore, early diagnosis of HCC is vital for patient survival, and the identification of biomarkers for early HCC detection is a crucial clinical need.[3]

Currently three FDA approved serum biomarkers are recommended to indicate the risk of liver cancer in high risk populations: AFP, AFP-L3, which is a fucosylated isoform of AFP, and des-gamma-carboxy-prothrombin (DCP).[4] However, these markers are not included in the surveillance guidelines published by the American Association for the Study of Liver Diseases[5] and the European Association for the Study of the Liver[6] because of low sensitivity and specificity. In China, AFP is used jointly with ultrasonography, computed tomography (CT), and pathology examination for early screening of HCC patients, as recommended by the Asian Pacific Association for the Study of the Liver.[7] This is based on the knowledge that the positive rate of AFP in HCC is about 60–80%.[8]

Better biomarkers are needed for HCC. The search for HCC biomarkers from clinical specimens using proteomics approaches mainly focused on blood samples in the past.[9,10] Besides, a few studies investigated biological fluids including urine[11] and tissue interstitial fluid.[12] To our knowledge, deregulated tissue proteins are presumably the main source of potential blood and bio-fluidic biomarkers, therefore, HCC tissue samples bear great potential in the HCC biomarker studies. Several groups have attempted to identify peptide ions by MALDI imaging MS to classify HCC tissues,[13,14] however, these studies did not characterize the respective proteomes. Li et al. studied malignant and matched benign tissues from 11 HCC patients using 2D LC coupled with shotgun proteomics with a gradient of 165 min LC and observed upregulation of SET complex proteins.[15] In another study, from 573 proteins identified by 2D gel electrophoresis and label-free shotgun proteomic analysis of seven pairs of HCC tissue samples, Megger et al. nominated 51 differentially expressed proteins.[16] In a more recent study, Naboulsi et al. identified and quantified 2736 proteins from 19 pairs of HCC tumor and adjacent benign tissues using the label-free shotgun proteomics over a 98-min LC gradient.[17] These studies based on shotgun proteomics and 2D gel electrophoresis analysis are generally of low throughput.

We have recently developed a methodology integrating PCT and sequential window acquisition of all theoretical mass spectra (SWATH-MS) for the rapid acquisition of proteotypes, defined as the acute quantitative state of a proteome, from biopsy-level tissue samples.[18] PCT technology produces high pressure (up to 45,000 psi) and lyses tissue samples by oscillating the pressure between low and high values, thereby allowing tissue lysis and protein digestion to occur in the same PCT-MicroTube. The method can be applied in a semi-automated and standardized manner and achieves relatively high sample throughput.[16–20] SWATH-MS is an emerging data-independent acquisition (DIA) mass spectrometric method that offers a high degree of quantitative accuracy, proteomic coverage, reproducibility of proteome coverage, and sample throughput.[21] Our previous proteomic study of prostate intra-tissue heterogeneity revealed most thus measured proteins from biopsy-level samples are quantitatively robust in prostate tissues.[22] The PCT–SWATH methodology was further demonstrated to be applicable to rapid proteotype acquisition from sub milligrams of tissue samples.[23] In our first study that demonstrated the technology, the PCT–SWATH method was applied to process and to convert 18 biopsy samples from nine patients with renal cell carcinoma into SWATH-MS fragment ion maps.[18] With a 2-h LC gradient, a 32 fixed SWATH windows setting, and the OpenSWATH data analysis tool,[21] about 2000 SwissProt proteins were identified with a high degree of reproducibility across all samples.[18] More recently, a more advanced and miniaturized device, i.e., PCT-MicroPestle,[19] was introduced providing for higher efficiency of protein extraction and peptide generation from biopsy-level tissues.

In this study, we first improved the SWATH methodology by adopting a 45-min LC gradient and an acquisition method using 67 variable SWATH windows to increase the sample throughput without compromising the proteomic depth. Then we applied the improved PCT–SWATH workflow to analyze peptide samples from a cohort of 19 HCC patients in technical duplicate with unprecedented reproducibility and throughput. In total, 11 787 proteotypic peptides from 2579 SwissProt proteins in 38 HCC tissue samples and 76 SWATH-MS runs were quantified. Sixteen proteins of high documented relevance to HCC are highlighted in our data with significant regulation in tumor tissues.

2. Experimental Section
2.1. Patients and Tissue Samples for SWATH-MS

Tissue specimens from HCC patients were collected from Union Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, China. All tissue samples were collected from hepatectomy specimens within 1 h after surgical removal, and were snap frozen and stored at −80 °C until proteomic analysis. Thirty-eight fresh frozen tissues from 19 HCC patients were included for PCT–SWATH analysis. For each patient, two tissue biopsy punches (with dimensions of 5 × 5 × 5 mm3) including a tumor tissue and a benign tissue from an adjacent region as determined by histomorphological
examination were collected. The clinical and pathological data for these patients are provided in Table S1, Supporting Information. The sample collection was approved by institutional review board of the Union Hospital of Wuhan in China. This study was approved by Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology.

2.2. PCT-Assisted Sample Preparation

For each sample, on average about 2 mg (weight range: 1.6–2.5 mg) of the frozen tissue was processed based on the protocol described previously,[18,19,24] Briefly, tissues were lysed with the PCT-MicroPestle device in 30 μL of lysis buffer composed of 8 M urea, 0.1 M ammonium bicarbonate, and protease inhibitor (Roche, Switzerland) in a barocycler HUB440 (Pressure BioSciences Inc, South Easton, MA, USA). The tissue lysis was performed under a program consisting of 60 oscillating cycles, with each cycle consisting 50 s of high pressure at 45,000 psi and 10 s of ambient pressure at 32 °C. Then the extracted proteins were reduced and alkylated by incubation with 10 mM tris (2-carboxyethyl) phosphine (TCEP) and 40 mM iodoacetamide (IAA) under gentle vortexing at 600 rpm for 30 min (ambient pressure), at 25 °C in the dark. Afterward, digestion was first performed with lys-C (Wako Japan; enzyme-to-substrate ratio, 1:40) in the barocycler (20,000 psi, 50 s high pressure and 10 s ambient pressure, 45 cycles), followed by trypsin (Promega, USA; enzyme-to-substrate ratio, 1:50) digestion in the barocycler (20,000 psi, 50 s high pressure and 10 s ambient pressure, 90 cycles). After digestion, the peptides were acidified with trifluoroacetic acid (TFA) to pH 2–3, and cleaned with SEP-PAK C18 (1 cc 50 mg) cartridges (Water, Milford, USA). The peptides were then dried under vacuum (CentriVap, Labconco, Kansas City, USA) and stored at -80 °C before MS analysis. Peptides were dissolved in HPLC-grade water containing 0.1% formic acid and 2% acetonitrile. Peptide concentration was measured with a NanoDrop 1000 spectrophotometer (Thermo Scientific, Wilmington, USA) via absorbance 280 nm (1 Ab = 1 mg·mL⁻¹).

2.3. SWATH-MS on 5600 TripleTOF

The peptide samples were spiked with iRT peptides (Biognos, Zurich, Switzerland) at a 1:10 (v/v) ratio. One microgram of peptide sample was injected to an Eksigent 1D+ Nano LC systems (Eksigent, Dublin) and analyzed in a 5600 TripleTOF mass spectrometer (SCIEX) in SWATH mode as described previously.[18] The LC gradient was reduced to 45 min (Figure S1, Supporting Information), and the SWATH acquisition scheme to 67 variable windows (Table S2, Supporting Information). The other SWATH parameters were set exactly as in our previous study,[18] except that the ion accumulation time for each SWATH window was 40 ms. Ion accumulation time for peptide precursors was set at 50 ms. The 38 samples were injected into the MS in randomized sequence once and then the same sequence was injected again to obtain a duplicate analysis. After each gradient, the column was washed twice using ramping gradient to minimize carryover (Figure S1, Supporting Information). Mass calibration using beta-gal was performed every fourth injections.

2.4. OpenSWATH Analysis

OpenSWATH (version 2014-12-01-154112)[25] was performed as described previously[18] except that a pan-human library (32 windows version)[23] was used to search the SWATH data in the iPortal platform (Table S3, Supporting Information, in our laboratory the entire pipeline was streamlined as an iPortal workflow).[26] The details of the pan-human library, containing 139,449 peptides from 10,316 proteins, have been described previously.[23] Briefly, SWATH miff files were converted into mzXML files using ProteinWizard mconvert.[27] The mzXML file was then converted to mzML file using OpenMS tool FileConverter.[28] OpenSWATH was performed using the tool OpenSWATHWorkflow with input files including the mzXML file, the TraML pan-human library file, and TraML file for iRT peptides.[29] Trypsin was selected as the proteotypic enzyme. The assays were identified with a false discovery rate below 0.1%. High confidence peptide features from different samples were aligned using the algorithm TRansition of Identification Confidence.[29] Label-free protein-level quantification was performed using the function ProteinInference from the R package aLFQ[30] on output from OpenSWATH.

2.5. Pathway and Process Enrichment Analysis by Metascape

Pathway and process enrichment analysis of the differentially abundant proteins were performed using Metascape (http://metascape.org). Terms with p-value < 0.01, minimum count 3, and enrichment factor > 1.5 (enrichment factor is the ratio between observed count and the count expected by chance) were collected and grouped into clusters based on their membership similarities. To further capture the relationship among terms, a subset of enriched terms was selected and rendered as a network plot, where terms with similarity > 0.3 were connected by edges.

Protein–protein interaction enrichment analysis was carried out for each given gene list (http://metascape.org). The resultant network contained the subset of proteins that form physical interactions with at least another list member. Molecular Complex Detection algorithm[31] was applied to identify densely connected network components. Pathway and biological process enrichment analysis was applied to each Molecular Complex Detection components independently and the three best-scoring (by p-value) terms were retained as the functional description of the corresponding components.

2.6. IHC Staining of HCC Tissues

For validation by IHC, three pairs of tissue sections (benign and tumor) from an independent HCC patient cohort, which was collected (Union Hospital, Wuhan, China, Table S4, Supporting Information) were cut out, fixed in formalin and stained using anti-MCM7 antibody (11225-1-AP, Proteintech China, Wuhan). This study was approved by Ethics Committee of Tongji Medical College, Huazhong University of Science and Technology.
2.7. Statistics

In volcano plots, paired Student’s t-test was used to compute p-values followed by Benjamini and Hochberg correction in R.

2.8. Data Deposition

All the SWATH Data files, library and original OpenSWATH results are deposited in PRIDE project[32]: PXD004873.

3. Results and Discussion

3.1. Identification of Peptides and Proteins by PCT–SWATH

To start, we extracted proteins from 38 biopsy-level liver tissue samples and produced tryptic peptides using the PCT method.[18,24] We measured the weight of each wet tissue sample and the total amount of peptide mass generated from the tissue and computed the peptide yield in weight percent per milligram of tissue. The results shown in Figure 1A indicate that the peptide yield for the 38 samples varied from 1.0% to 9.5%, with an average value of 6.0%. A total of 59.6 μg of total peptide mass per milligram wet liver tissue was obtained on average. The peptide yield in tumor tissue (4.9% on average) was lower than that of benign tissue (8.1% on average, p-value = 0.00037) (Figure 1A). The results further showed that tumor samples exhibited more variability in terms of peptide yield compared to the benign tissue samples. The CV of peptide yield was 15.8% for benign tissue samples (n = 19), 40.7% for tumor samples (n = 19), and 32.9% overall (n = 38) (Figure 1B), demonstrating the heterogeneity of liver tumors.

Each peptide sample was then analyzed using SWATH-MS in technical duplicate, using a 45-min LC gradient and a 67 variable SWATH window acquisition scheme (Table S2, Supporting Information), which was optimized based on the ion density map of human proteins extracted from HCC tissues. With this SWATH window setting, up to 20 samples were analyzed per MS instrument per 24 h. Mass calibration using beta-gal was performed every four injections. The MS analyses of the 76 SWATH runs were completed in about 4 working days.

After searching the resulting raw data files against the pan-human SWATH assay library,[23] 2579 unique proteotypic Swiss-Prot proteins with 11 787 unique peptides were identified at a false discovery rate (FDR) below 1% using OpenSWATH[23] (Table S5, Supporting Information). Comparable numbers of proteins and peptides were quantified from both benign and tumor tissue samples (Figure 2A and B). In this study, the PCT–SWATH throughput was further increased compared to the previous work.[18] The protein numbers identified per sample were slightly higher than those obtained from human kidney tissues.[18]

3.2. Variability of Protein Expression between Tumor Tissue and its Adjacent Benign Pair

To evaluate the quantitative accuracy of the method, we computed the global Pearson correlation of the two technical replicates at the peptide level for all samples after removing missing values. The R² was 0.9729 (Figure 2C), indicating that our label-free method could quantify peptides with high degree of reproducibility. The global view of protein quantification is shown in Figure 2D. It showed that all the technical replicates were clustered together respectively.

We then analyzed the overall variation of protein abundance for all 2579 quantified proteins from tumor tissues, adjacent benign tissues, and all samples, respectively (Figure 3A). The data showed that the CV of tumor tissue samples was slightly greater than that of adjacent benign tissues, indicating more overall variability of protein expression in tumors. We next checked the relationship between protein abundance and the overall variation among these samples and observed that low abundance proteins exhibited a higher degree of overall variability (Figure 3B–D). In Figure 2C, proteins of low abundance demonstrated comparable reproducibility with high abundance proteins, therefore, we mainly attributed the higher variability in tumor tissue to biological causes.

3.3. Differentially Regulated Proteins Characterized in the HCC Cohort

We further explored differentially expressed proteins between tumor and benign samples. A total of 541 proteins showed significant differential abundance (Figure 4A). Of these, 381 proteins showed increased abundance and 160 proteins showed decreased abundance in tumor tissues compared to benign tissues. Pathway enrichment analysis of the regulated proteins in HCC tumor tissues was performed by Metascape. In total, 24 pathways were distinguished to be activated and 13 pathways were found to be suppressed, as shown in Table S6, Supporting Information. We found that proteins with increased abundance were highly enriched in mRNA related processes including mRNA metabolic process, peptide metabolic process, ribonucleoprotein complex biogenesis, and nuclear transport. The proteins with decreased abundance were found to be mostly involved in small molecule metabolic processes such as small molecule catabolic process, monocarboxylic acid metabolic process, and carbon metabolism. Next, we carried out protein–protein interaction network analysis using Metascape. The resultant network contains the subset of proteins that form physical interactions with at least another protein showing altered abundance.

Heat shock proteins (HSP) were found to be upregulated in our dataset, such as HSPA1A, HSPA1L, HSPA4, HSPA8, HSPB1, and HSP90AB1 (Figure 4B, Table S6, Supporting Information). HSP proteins constitute a family of highly conserved stress response proteins that can protect cells and induce them to repair damage caused by a variety of stress.[33] Several members of the HSP family have been reported to be overexpressed in a wide range of human cancers.[34] It has been demonstrated that 14-3-3σ protein induces HSP70 via a β-catenin/HSF-1-dependent pathway, which consequentially modulates HCC. The HSF-1/Hsp70-1 protein complex might be of prognostic value for HCC.[35] We did not detect HSP70 in this dataset, but a series of heat shock proteins have been characterized to be overexpressed in this cohort.
Figure 1. Performance of PCT-assisted peptide preparation. A) Peptide yield in two types of liver tissue samples. Y-axis shows the peptide yield in percentage of peptide (μg) per milligram (mg) tissue. (B) CV of peptide yield in two types of tissues and across all samples.

We also observed that proteins related to mitochondrial translation, specifically the mitochondrial ribosomal proteins (MRPs), were among the most strongly upregulated protein clusters (Figure 4C). Tumor initiation and progression in cancer cells involve the development of mechanisms to inhibit apoptosis at multiple stages, whereas mitochondria play crucial roles in the induction of apoptosis and programmed cell death. This paradigm is central to malignant cellular transformation because altered mitochondrial function and defective apoptosis are well-known hallmarks of cancer cells. It has also been reported that the expression of genes encoding MRPs, mitoribosome assembly factors, and mitochondrial translation factors is modulated in multiple cancers, which is linked to tumorigenesis and metastasis.\(^{[36]}\) Mini-chromosome maintenance complex component 7 (MCM7) was also found to be overexpressed (Figure 4C). MCM7 plays an essential role in initiating DNA replication. DNA replication is a central process in cell proliferation, while aberrant DNA replication is indicated to be a driving force of oncogenesis. It has been reported that overexpressed MCM7 is associated with a poor prognosis of HCC patients, and MCM7 promotes HCC cell proliferation via upregulating MAPK-cyclin D1 pathway both in vitro and in vivo.\(^{[17]}\) In addition, the hallmark MYC targets proteins, and the cell cycle checkpoint proteins involved in G2/M transitions, are also among the upregulated protein networks.

Small molecule catabolic processes characterized the group of proteins that showed the strongest reduction of abundance in the HCC samples tested. Our data showed that proteins involved in glycogen metabolism pathway were suppressed in HCC samples (Figure 4D). These proteins include the glycogen synthesis enzymes including phosphoglucomutase-1 (PGM1), and UTP-glucose-1-phosphate uridylyltransferase (UGP2), as well as the glycogen degradation enzyme liver glycogen phosphorylase (PYGL). They are downregulated in tumor tissues compared to their matched benign controls, indicating the reprogramming of glycogen metabolism in HCC, which suppresses the turnover of glycogen. Glycogen metabolism is a key pathway induced by hypoxia, and is necessary for optimal glucose utilization. In mammalian cells, glycogen is the main storage form of glucose. The liver accumulates glycogen in large amounts and releases it slowly to maintain blood glucose levels.\(^{[38]}\) Recently glycogen metabolism has become a recognized feature of cancer cells. As to HCC, it has been reported that inhibition of glycogen synthase kinase-3β (GSK-3β) suppressed the growth and induced caspase-dependent apoptosis in HCC cells, and GSK-3β may become a promising therapeutic target for HCC.\(^{[39–41]}\) In our dataset, glycogen metabolism pathway was found to be suppressed in HCC patients. Besides, our data did not indicate whether this reprogramming changed the direction of glycogen storage.

The ethanol metabolism was also found to be downregulated. Aldehyde dehydrogenase (ALDH) is a gene superfamily that is responsible for the detoxification of biogenic and xenogenic aldehydes. Both ALDH2 and ALDH1B1 proteins were significantly downregulated in our dataset (Figure 4D). Similarly, a previous study by Park et al. revealed increased detection of ALDH-3 but decreased detection of ALDH-2 when measured by 2D electrophoresis and MS in a group of ten cases with HCC.\(^{[42]}\) ALDH1 and ALDH2 are the most important enzymes for aldehyde oxidation. The ethanol detoxifying pathway in humans occurs mainly in the liver and is carried out by two enzymatic steps. In the first step, ethanol is metabolized quickly by alcohol dehydrogenase (ADH) to generate acetaldehyde, and the latter is then metabolized by the mitochondrial ALDH2 to acetate.\(^{[43]}\) ALDH2 is expressed ubiquitously in all tissues but is most abundant in the liver and found in high amounts in organs that require high mitochondrial oxidative phosphorylation. The downregulation of ALDH2 in liver tumor tissue will likely accumulate toxin acetaldehyde and promote tumorigenesis.

In summary, liver is the pivotal organ for molecular synthesis and the metabolic hub of human beings. Our data showed that the molecular synthesis related pathways were enhanced, while most major metabolic processes were suppressed in the HCC tissues tested, suggesting a biochemical imbalance in HCC cells.
Figure 2. Peptide and protein identification and quantification. A) Number of peptides identified in each sample (76 SWATH runs of 38 peptide samples from 19 pairs of HCC patient tissues). B) Number of proteins identified in each sample. C) Global Pearson correlation of technical replicates of peptide quantification for samples in all samples. Values are log2-transformed protein intensity. Missing values were excluded from this analysis. D) Heatmap overview of protein abundance patterns in 76 samples. Protein intensity values were log10 transformed and subjected to unsupervised-clustering in two axes. (P, patient ID; N, benign tissue; T, tumor tissue; a, the first technical replicate; b, the second technical replicate; for example, P1Na, the technical replicate 'a' of benign tissue from patient no. 1.)

3.4. Proteins with Altered Abundance in HCC Patients with High Serum AFP

Next, we explored the molecular pathogenesis of HCC by associating protein expression with various clinic-pathological characteristics of HCC. AFP is the most widely used tumor biomarker currently available for early detection of HCC, and the widely accepted threshold of serum AFP is 20 ng mL⁻¹. It has been reported that at this threshold, serum AFP had a sensitivity of 41–65% and specificity of 80–94%.[44] We grouped the 19 HCC patients in our cohort into two groups, according to an AFP-cutoff of 20 ng mL⁻¹ (Table S1, Supporting Information), computed the fold change of protein expression between tumor and benign tissues, and checked the protein regulation pattern of 11 patients with high AFP level (>20 ng mL⁻¹) and six patients with low AFP level (<20 ng mL⁻¹), respectively. Two patients without serum AFP examination value at surgical operation were excluded from these analyses. Although a considerable proportion of HCC patients in the cohort do not have elevated serum AFP, and serum AFP can increase in patients with diseases other than HCC,[4] high expression of AFP in serum correlated with high cell proliferation, high angiogenesis, and low apoptosis and was associated with poor prognosis of HCC.[45] Indeed, we found in our data set that the low serum AFP level was closely related to higher survival rate of the patients (five out of six patients have recurrence free survival), especially when there was no portal vein thrombosis detected by CT at diagnosis, as is shown in Table S1, Supporting Information. On the other hand, in the AFP-high patient group, six of 11 patients were dead and two patients suffered from tumor recurrence, displaying much poorer survival (Table S1, Supporting Information). Nevertheless, statistical analysis was not feasible due to small sample size.
With respect to the 11 AFP-high patient group, 419 proteins showed increased abundance and 192 proteins showed decreased abundance (p-value < 0.05) in tumor tissues compared to tumor tissues (Figure 4E). In the AFP-low subgroup of six patients, our data showed that no protein significantly regulated between the tumor and benign tissues, probably due to the small sample size (Figure 4F). Pathway enrichment and network analysis of the differentially expressed proteins from the AFP-high subgroup revealed similar regulation pattern with the 19 HCC patients group.

In the AFP-high subgroup, we identified additional downregulated proteins including methylene tetrahydrofolate dehydrogenase 1 (MTHFD1) (Figure 4G), which is an enzyme identified in the one-carbon (1C) metabolism pathway. Another enzyme in the pathway, i.e., serine hydroxymethyltransferase (SHMT1), was also downregulated (Figure 4D). The data suggest reprogramming of one-carbon metabolism in HCC. This is consistent with the Liverome study reporting that downregulation of MTHFD1 is prevalent in proliferative and poorly differentiated HCCs[46].

We also observed downregulation of a tumor suppressor protein regucalcin (RGN) (Figure 4H), which modulates multiple protein kinases and phosphatases in cancers including HCC.[47] Transcripts analysis has reported that prolonged survival in HCC patients is associated with increased RGN gene expression.[48] Our data at protein level further consolidate the potential roles of RGN in HCC.

Two proteins involved in lipid metabolism showed decreased expression in the AFP-high HCC tissues. They are ACA2 and ANXA6 (Figure 4H). ACA2 is a mitochondrial enzyme involved in the fatty acid metabolism pathway. ANXA6 is a lipid-binding protein highly expressed in the liver, regulating cholesterol homeostasis, and signaling pathways with a role in liver physiology. A recent study found that ANXA6 was downregulated in HCC tissues compared to benign tissues in 18 HCC patients.[49] The data indicate dysfunction of lipid metabolism in HCC tissues.

3.5. Quantitative Proteomics Analysis of Hepatitis B Virus Positive HCC Subgroup

Chronic hepatitis B virus (HBV) is the most common cause of HCC in China. Approximately 80% of HCC develops from liver cirrhosis, which predominantly progresses from HBV in China.[50] We found that 106 and 75 proteins showed increased or decreased abundance, respectively, in 11 HCC patients with HBV infection (Figure 4I, Table S6, Supporting Information) compared to those without HBV infection. Subsequent pathway and protein interaction analysis by Metascape revealed a number of interesting protein clusters that showed a close relationship to liver cirrhosis in patients with chronic HBV infection. Pyruvate kinase M (PKM) was found to be upregulated in this HBV
Figure 4. Significantly regulated proteins in tumor versus benign tissues from the 19 HCC patient cohort (ALL), the high blood AFP level subgroup (AFP\_High) of 11 HCC patients, the low blood AFP level subgroup (AFP\_Low) of six HCC patients, and the HBV infected subgroup (HBV\+) of 11 HCC patients, respectively. A) Volcano plot of proteins expressed in tumor tissue samples and their matched benign samples from the “ALL” group. Each data point represents the ratio of tumor to adjacent benign tissue observed for a protein. The horizontal line at $-\log_{10}(p) = 1.3$ indicates $p = 0.05$. Data points above the line ($p < 0.05$) are considered to be significantly regulated. The two vertical lines at $\log_2(T/N) = \pm 1$ where the fold change is equivalent to 2 indicate the threshold for up-/downregulation. The upper left quadrant ($p < 0.05$ and fold change $< 0.5$) represent the downregulated proteins whereas the upper right quadrant indicates the upregulated proteins in tumor tissue. N, benign; T, tumor. B and C) Enriched significantly upregulated protein networks for the 19 HCC patient cohort. D) Enriched significantly downregulated protein network for the 19 HCC patient cohort. E) Volcano plots of proteins expressed in tumor tissue samples and their matched benign samples in the 11 AFP-high patients group. F) Volcano plot of proteins expressed in tumor tissue samples and their matched benign samples in the six AFP-low patients group. G and H) The enriched significantly downregulated protein network in the AFP-high patient group. I) Volcano plot of proteins expressed in tumor tissue samples and their matched benign samples in 11 HBV\+ HCC patients. J) The enriched significantly upregulated protein networks in the HBV positive patient group. K and L) The downregulated protein networks in the HBV positive patient group.
Figure 5. Summary of regulated proteins and pathways in HCC tumor tissues compared to matched benign samples. A) All proteins found regulated in HCC tissues, including AFP-high group and HBV\(^+\) group, were inspected in literature. Regulated pathways and their component proteins are displayed. B) 16 selected proteins significantly regulated in HCC tumor tissues compared to matched benign samples. For each protein, its expression, after log\(_{10}\) transformation, in benign (N) and tumor (T) tissues are shown in boxplots. Student’s t-test was used to compute the significance between N and T samples.

positive (HBV\(^+\)) subgroup (Figure 4J). PKM is a protein kinase that catalyzes the final step in glycolysis by transferring the phosphate from phosphoenolpyruvate (PEP) to ADP, thereby generating pyruvate and ATP. PKM has two isoforms, PKM1 and PKM2. PKM2 plays an important role in metabolic alterations related to inflammation and cancer. The ATP generation by PKM2, unlike mitochondrial respiration, is independent of oxygen and thus allows tumor cells to grow in hypoxic conditions. Cancer cells are characterized by high glycolytic rates to support energy regeneration and anabolic metabolism, along with the high expression of PKM2. It has been shown in a recent report that the anti-apoptotic protein poly (ADP-ribose) polymerase (PARP)14 promotes aerobic glycolysis in HCC by maintaining low activity of PKM2, and that the PARP14-JNK1-PKM2 regulatory axis links apoptosis and metabolism. In agreement with our finding that HBV\(^+\) tumors expressed higher level of PKM, there is another report showing that high PKM2 expression was more frequently found in cirrhotic liver caused by HBV infection than...
in non-cirrhotic liver, and that PKM2 overexpression was associated with poor survival rates in HCC patients with cirrhotic liver (CL).[^54]

Cytidine triphosphate synthase 1 (CTPS1) is another example of a protein that showed higher expression in the HBV+ HCC subgroup in this study compared to patients without HBV infection (Figure 4J). CTPS1 catalyzes the rate-limiting step in the de novo CTP synthetic pathway, in which a UTP is converted into CTP with the consumption of glutamine and ATP. In fact, upregulated CTPS1 activity has been observed in multiple human and rodent tumors,[^55] promoting tumor transformation and progression.[^56] Recently, researchers focus on the CTPS cytoophidium that is a filamentous intracellular macrostructure aggregated by CTPS, and is able to promote the activity of CTPS.[^57] A recent study examined the presence of CTPS cytoophidia in tumor and the adjacent benign tissues of HCC patients, and found that many cytoophidia were observed in 28% of the HCC tumor samples but not the adjacent hepatocyte population.[^58] In addition, they found that the high expression of HSP90 was correlated with the presence of CTPS cytoophidia, which is consistent to our findings that heat shock proteins (HSPA1L and HSPA8) and CTPS1 are all upregulated (Figure 4J).

On the other hand, both cystathionine β-synthase (CBS) and cystathionine-γ-lyase (CTH) were found at decreased abundance in the HCC samples tested (Figure 4K). CBS and
CTH are abundant proteins in the liver, endogenously produce hydrogen sulfide (H$_2$S), a gasotransmitter modulating synaptic transmission, vasorelaxation, angiogenesis, inflammation, and cellular bioenergetics. Regulation of H$_2$S influences lipid metabolism, glucose metabolism, oxidative stress, and mitochondrial bioenergetics.[19] Suppression of CBS transcripts has been reported in a study of 120 HCC patients compared to noncancerous liver tissue.[60] Reduced CBS transcripts expression is significantly correlated with high tumor stage, high Edmondson grade, and high AFP level.[60] However, neither the CBS nor the CTH protein level has yet been reported to be investigated in HCC tissue. Our data suggest downregulation of H$_2$S production via suppressing CBS and CTH in HCC tissues.

Fructose-bisphosphate aldolase B (ALDOB) is an enzyme for glucose and fructose metabolism. In a cohort of 313 HCC patients, ALDOB was found downregulated in HCC tissue using tissue microarray and IHC technologies, and its downregulation was significantly correlated with HCC progression.[65] Furthermore, ALDOB has been found to be a binding protein of the S region of the HbsAg.[66] Our data showed that ALDOB expression was suppressed in HBV$^+$ HCC patients (Figure 4L), suggesting a potentially interesting mechanistic linkage between HCC and HBV infection, which might be utilized as therapeutic target.

3.6. Summary of Regulated Proteins in HCC Samples

We summarized all differentially expressed proteins in HCC samples including the AFP-high group and HBV$^+$ group (Figure 5A). Interestingly, the proteins with increased expression concentrated in three types of pathways: a) production of DNA, mRNA and protein; b) oncogenic signaling pathways and apoptosis; c) immune response. This suggests enhanced proliferation of tumor cells in the malignant tissues. In contrast, proteins participated in various metabolism pathways were downregulated, indicating the suppression of metabolic functionality in liver during oncogenesis.

From literature review of all these regulated proteins in Figure 5A, we identified 16 proteins of reported clinical relevance to HCC (Figure 5B). We calculated the expression level of these 16 proteins between tumor and the adjacent benign tissue, as was shown in Figure 5B, the log$_{10}$ MS intensity, with the p-value for each pair of comparison between benign and tumor.

Among these potential candidates, MTHFD1, RGN, CBS, and CTH were previously reported to be dysregulated in HCC as transcripts at gene level, as discussed above. The protein expression level of MCM7, ALDH2, ANXA6, PKM2, CTPS, and ALDOB were reported to be changed in HCC in other studies. The others have not been reported yet.

3.7. Detection of MCM7 in HCC Samples Using IHC

To verify whether findings from our SWATH data are consistent with IHC reports, we further applied MCM7 IHC staining to three additional patients who accepted surgical operation very recently (Table S4, Supporting Information) and observed the up-regulation of MCM7 in liver tumors (Figure 6). The data collectively suggest that upregulation of MCM7 might play crucial roles in liver cancer progression.

4. Conclusions

In conclusion, here we report an optimized PCT–SWATH workflow enabling analysis of clinical tissue specimens with increased sample throughput without compromise of quantitative accuracy and proteomic coverage. Our study identified a few regulated proteins in this HCC cohort. Proteins with increased abundance are mainly related to mass production, oncogenic signaling, and immunity, whereas metabolic proteins are shown with lower expression. We identified 16 dysregulated proteins of reported clinical relevance to HCC. Upregulation of MCM7 in tumor tissue samples was observed using IHC in additional patients. The study demonstrates that the PCT–SWATH methodology has potential to be practically applied in clinical research to analyze tissue samples in high throughput.

Supporting Information

Supporting Information is available from the Wiley Online Library or from the author.

Acknowledgements

Y.Z. and J.Z. are the co-first authors of this study. T.G. conceived and coordinated the project. P.S. procured the tissue samples and helped with finding the relevant matching tumor/benign sample pairs. J.Z. and C.L. punched all tissue blocks. J.Z., C.L., W.X., and Y.Z. performed the PCT-based sample preparation. Y.Z. and T.G. performed the SWATH-MS data acquisition. P.S., J.Z., and C.L. provided clinical follow-up data. X.D. supervised Y.S. and X.Y. to evaluate the MCM7 IHC staining pictures. J.Z., Y.Z., Q.Z., L.Y., T.Z., and T.G. performed the data analysis. Y.Z., J.Z., R.A., S.H., and T.G. wrote the manuscript with inputs from all co-authors. R.A., S.H., and T.G. supervised the project. This study was supported by the National Natural Science Foundation, P.R China (NNSF/81200348/2013), the Swiss National Science Foundation (SNF 166435 MitoModules to R.A.) ERC PROTEOMICS4D (project no. 670821) and Westlake University Startup funds to T.G. Y.Z. was supported by a fellowship from SignalX (SystemsX RTD 2013/156) and EU PeCISE (project no. 668858).

Conflict of Interest

R.A. holds shares of Biognosys AG, which operates in the field covered by the article. The research groups of R.A. and T.G. are supported by AB SCIEX, which provides access to prototype instrumentation, and Pressure Biosciences, which provides access to advanced sample preparation instrumentation.

Keywords

biopsy tissue, hepatocellular carcinoma, pressure cycling technology (PCT), SWATH

[1] H. B. El-Serag, K. L. Rudolph, Gastroenterology 2007, 132, 2557.
[60] J. Kim, S. J. Hong, J. H. Park, S. Y. Park, S. W. Kim, E. Y. Cho, I. G. Do, J. W. Joh, D. S. Kim, Oncol Rep, 2009, 21, 1449.

[61] Q. F. Tao, S-x. Yuan, F. Yang, S. Yang, Y. Yang, J-h. Yuan, Z-g. Wang, Q-g. Xu, K-y. Lin, J. Cai, J. Yu, W-l. Huang, X-l. Teng, C-c. Zhou, F. Wang, S-h. Sun, W-p. Zhou, Mol Cancer, 2015, 14, 170.

[62] J. Wu, C. Dan, H. B. Zhao, C. X. Xiao, Y. P. Liu, L. J. Si, J. L. Ren, B. Guleng, Crit Rev Eukaryot Gene Expr, 2014, 24, 181.