In-depth proteomic analysis of tissue interstitial fluid for hepatocellular carcinoma serum biomarker discovery

Jian Zhang[1], Ning Hao[1], Wei Liu[2], Min Lu[3], Longqin Sun[1], Ning Chen[1], Miantao Wu[4], Xiaohang Zhao[5], Baocai Xing[2,7], Wei Sun*[,1,7] and Fuchu He*[,1,6,7]

[1]State Key Laboratory of Proteomics, Beijing Proteome Research Center, National Center for Protein Sciences Beijing, Beijing Institute of Radiation Medicine, Beijing 102206, China; [2]Key Laboratory of Carcinogenesis and Translational Research (Ministry of Education/Beijing), Department of Hepato-Pancreato-Biliary Surgery I, Peking University Cancer Hospital and Institute, Beijing 100036, China; [3]Department of Pathology, School of Basic Medical Sciences, Peking University, Beijing 100191, China; [4]State Key Laboratory of Oncology in South China, Collaborative Innovation Center for Cancer Medicine, Sun Yat-sen University Cancer Center, Guangzhou 510060, China; [5]State Key Laboratory of Molecular Oncology, Cancer Institute and Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing 100021, China and [6]Institutes of Biomedical Sciences, Fudan University, Shanghai 200032, China

Background: Hepatocellular carcinoma (HCC) is a primary malignancy of the liver. New serum biomarkers for HCC screening are needed, especially for alpha-fetoprotein (AFP) negative patients. As a proximal fluid between body fluids and intracellular fluid, tissue interstitial fluid (TIF) is a suitable source for serum biomarker discovery.

Methods: Sixteen paired TIF samples from HCC tumour and adjacent non-tumour tissues were analysed by isobaric tags for relative and absolute quantitation (iTRAQ) method. Two proteins were selected for ELISA validation in serum samples.

Results: Totally, 3629 proteins were identified and 3357 proteins were quantified in TIF samples. Among them, 232 proteins were significantly upregulated in HCC-TIF and 257 proteins down-regulated. Two overexpressed extracellular matrix proteins, SPARC and thrombospondin-2 (THBS2) were selected for further validation. ELISA result showed that the serum levels of SPARC and THBS2 in HCC patients were both significantly higher than those in healthy controls. The combination of serum SPARC and THBS2 could distinguish HCC (AUC = 0.97, sensitivity = 86%, specificity = 100%) or AFP-negative HCC (AUC = 0.95, sensitivity = 91%, specificity = 93%) from healthy controls. And the combination of serum SPARC and THBS2 could also distinguish HCC patients from benign liver disease patients (AUC = 0.93, sensitivity = 80%, specificity = 94%). In addition, serum THBS2 was found to be a novel independent indicator for poor prognosis of HCC.

Conclusions: Novel HCC candidate serum markers were found through in-depth proteomic analysis of TIF, which demonstrated the successful utility of TIF in cancer serum biomarker discovery.

Hepatocellular carcinoma (HCC) is one of the most common malignancies worldwide. The mortality of HCC is high mainly due to the late diagnosis. Screening programs have improved the early detection of HCC (Sherman, 2005), especially in high-incidence regions such as Asia (Mittal and El-Serag, 2013). Alpha-fetoprotein (AFP) is recommended for HCC screening and surveillance by many guidelines from Asia such as the Asian Pacific Association for the Study of the Liver (APASL) (Omata *et al*, 2010) and the

Japan Society of Hepatology (JSH) (2010). The elevation of AFP is a risk factor for HCC and can be used to help define patients at risk. However, the sensitivity of AFP is insufficient (Daniele et al, 2004). Thus, new biomarkers are urgently needed, especially for AFP-negative HCC patients.

As a kind of body fluid commonly used in the clinic, blood is an ideal source for biomarker discovery. However, identification of low-abundance proteins in the blood remains a challenge due to the complexity of the blood proteome. Tissue-based proteomic studies have also been carried out for biomarker discovery (Kikuchi et al, 2012). But only secreted proteins with dysregulated abundance in blood have the potential for clinical application. How to get the tumour-originated proteins which could be released into the blood is an important issue in cancer biomarker discovery.

In this regard, proximal fluids have emerged as a novel source. Tissue interstitial fluid (TIF), as the proximal liquid surrounding the tissue cells, could reflect the microenvironment alterations during tumour development. Compared with serum, TIF contains less high abundant proteins and more suitable for mass spectrometry analysis (Sun et al, 2010). Besides, the concentration of proteins secreted or shed from the tumour cells is higher in TIF than that in serum. Those characteristics make TIF a good source for biomarker discovery. In fact, many groups have proved that TIF could be used for biomarker discovery since Celis' group (Celis et al, 2004) established tissue elution method for TIF preparation. For example, Teng et al (2011) analysed the TIF of renal cell carcinoma (RCC) and validated the elevation of enolase 2 in the serum of RCC patients. Xie et al (2016) identified six serine proteases in colon cancer TIF and four of them were validated in serum.

Our group (Sun et al, 2010) has established the extraction method of liver TIF in previous work. In this study, in order to identify serum markers for AFP-negative HCC patients, paired tumour and adjacent non-tumour TIFs from 16 HCC patients with low serum AFP levels were analysed with four batches of 8-plex iTRAQ. Bioinformatics analysis showed that proteins involved in extracellular matrix (ECM) organisation were significantly enriched in HCC TIF and in one subtype of HCC patients with poor prognosis. The ECM is a collection of extracellular molecules, which composed primarily of collagens, laminins and various glycoproteins. The ECM proteins, which could regulate inter-cellular communication, may play an important role in promoting tumour growth and metastasis. Two ECM-associated proteins, SPARC and thrombospondin-2 (THBS2), were selected for further validation. SPARC is a secreted protein which may support tumour-stroma interactions and contribute to cancer progression in some certain cancers such as lung cancer (Wong and Sukkar, 2017). THBS2 is a secreted glycoprotein which may have an oncogenic function in lung adenocarcinoma and breast cancer but may have an anti-angiogenic role in gastric cancer (Weng et al, 2016). Our study found that SPARC and THBS2 could be candidate serum screening markers for HCC, including AFP-negative HCC. THBS2 was also identified to be a candidate serum prognostic marker for HCC.

MATERIALS AND METHODS

Sample preparation. Paired tumour and non-tumour liver tissues from 16 HCC patients who underwent hepatectomy were collected for TIF extraction. All the patients were with serum AFP $< 100 \, \mathrm{ng \, ml^{-1}}$, including 12 patients with serum AFP $< 20 \, \mathrm{ng \, ml^{-1}}$. The clinical characteristics of the patients (Supplementary Table S1) were summarised in Table 1. Blood samples from 44 HCC patients, 17 benign liver disease patients and 30 healthy donators were collected for ELISA analysis

Table 1. Clinical characteristics of HCC patients involved in iTRAQ experiment

Clinical characteristics	HCC group ($N = 16$) n (%)
Age (year)	
< 60	13 (81.2%)
$\geqslant 60$	3 (18.8%)
Gender	
Male	14 (87.5%)
Female	2 (12.5%)
HBV	
Positive	8 (50.0%)
Negative	8 (50.0%)
Tumour size (cm)	
< 5	6 (37.5%)
$\geqslant 5$	10 (62.5%)
Microvascular invasion	
Positive	1 (7.7%)
Negative	13 (92.3%)
AFP (ng ml^{-1})	
< 20	12 (75.0%)
20–100	4 (25.0%)
TNM stage	
I	7 (43.8%)
II	4 (25.0%)
III	5 (31.2%)
Abbreviation: HCC = hepatocellular carcinoma.	

(Supplementary Table S2). Blood samples were centrifuged at 2000 g for 5 min at 4 °C within 1 h of collection. The supernatant was further centrifuged at 13 500 g for 15 min. The serum was aliquot and stored at −80 °C until use. HCC samples were from Peking University Cancer Hospital. Benign liver disease samples were from Peking University Cancer Hospital and Sun Yat-sen University Cancer Center. Healthy samples were from Cancer Hospital of Chinese Academy of Medical Sciences. All patients provided informed consents. The use of human samples was approved by the Ethics Committees of Peking University Cancer Hospital, Sun Yat-sen University Cancer Center and Cancer Hospital of Chinese Academy of Medical Sciences.

TIF collection. The TIF was collected according to the protocol described before (Sun et al, 2010). In brief, fresh tissues were cut into 1–3 mm^3 pieces and washed carefully with PBS. After incubation in a CO_2 incubator at 37 °C for 2 h, the samples were centrifuged at 1000 g for 3 min to remove cells and the supernatant was transferred to a new tube for further centrifugation at 2000 g for 8 min and then 20 000 g for 30 min at 4 °C. TIF samples were aliquot and stored at −80 °C. The protein concentration of TIF was measured using the Bradford assay (Bio-Rad Laboratories, Hercules, CA, USA).

iTRAQ analysis of TIF samples. For each TIF sample, 50 μg of proteins were mixed with 200 μl of UA buffer (8 mol l^{-1} urea in 0.1 mol l^{-1} Tris/HCl pH 8.5) in Amicon Ultra-0.5 Centrifugal Filter (10 kDa; Millipore Corp, Billerica, MA, USA) and centrifuged at 14 000 g for 20 min. Then samples were treated with 20 mmol l^{-1} dithiothreitol at 37 °C for 4 h. After centrifugation, 50 mmol l^{-1} iodoacetamide was added and incubated for 30 min at room temperature in the dark. After washing with UA buffer, triethylammonium bicarbonate buffer (TEAB) was added and the sample was digested with trypsin (Promega, Madison, WI, USA; enzyme to protein ratio 1 : 50) at 37 °C overnight. Digested peptides were labelled with iTRAQ reagents (Applied Biosystems, Foster City, CA, USA) according to the manufacturer's

instructions. For each 8-plex iTRAQ, tumour TIF samples were labelled with iTRAQ tags 113, 115, 117, 119, and paired adjacent non-tumour TIFs were labelled with iTRAQ tags 114, 116, 118, and 121, respectively. To determine the ratio threshold for differential proteins, the pooled tumour and adjacent non-tumour TIF samples were analysed twice (labelled with iTRAQ tags 113 and 114, or 115 and 116, respectively). Equal amounts of iTRAQ-labelled peptides were mixed and dried, then resuspended in buffer A (2% acetonitrile, 98% water with ammonia at pH 10) and fractionated by HPLC (RIGOL, Beijing, China). For each experiment, 11 fractions were collected and dried. Then peptides were redissolved with 0.1% formic acid (FA) and analysed on a Q-Exactive HF mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA) coupled with a nano-HPLC instrument (UltiMate 3000 LC system, Dionex, CA, USA). The mass spectrometer was operated in the data-dependent mode with positive polarity at electrospray voltage of 2 kV. Full scan MS spectra (m/z 300–1400) were acquired followed by MS/MS on the top 20 intense ions detected. The mass spectrometry (MS) raw data were analysed with Proteome Discoverer software (version 1.4) using the Mascot search engine to search against the human database (UniProtKB, release 2014_09). The following parameters were applied: precursor mass tolerance was 15 p.p.m.; fragment tolerance was 20 m.m.u.; the dynamic modifications were oxidation (M) and iTRAQ labelling (K, Y, and N-term); the static modification was carbamidomethyl (C); a maximum of two missed cleavages were allowed. Peptides with peptide score $\geqslant 10$ and FDR < 0.01 (based on the target-decoy database algorithm) were used for protein grouping. Protein groups identified $\geqslant 2$ peptides from all samples were considered for further analysis and only unique peptides were used for protein quantification. Gene Ontology (GO) enrichment analysis was performed with DAVID 6.8 (Huang da *et al*, 2009). K-means consensus clustering was performed with R package ConsensusClusterPlus. Only proteins quantified in all samples and s.d. of log2 (ratio$_{tumour/non-tumour}$) $\geqslant 0.7$ were selected for consensus clustering (Mertins *et al*, 2016). After normalisation with quantile using R package limma, consensus clustering was performed using k-means as cluster algorithm and euclidean as distance algorithm with up to 6 clusters. Subtype signatures were selected based on proteins with $< 50\%$ missing data by ANOVA test and one-*vs*-rest *t*-test (Tyanova *et al*, 2016a).

Western blot. TIF proteins were electrophoresed on 10% SDS–PAGE gels and transferred to nitrocellulose membranes. After being blocked for 1 h at room temperature in TBST (0.05% Tween-20 in 20 mol l^{-1} Tris-HCl, 140 mol l^{-1} NaCl, pH 7.5) containing 5% skim milk, membranes were incubated overnight at 4 °C with goat anti-SPARC (AF941; R&D Systems, Minneapolis, MN, USA) or goat anti-THBS2 (sc-7655; Santa Cruz Biotechnology, Santa Cruz, CA, USA) antibodies. GAPDH (CW0100A; CWBIO, Beijing, China) was used as an internal reference. After four times of washes with TBST, the membranes were incubated for 1 h at room temperature with horseradish peroxidase (HRP)-conjugated secondary antibodies. After another four times of washes with TBST, the membranes were detected with ECL reagent (Thermo Fisher Scientific, Waltham, MA, USA).

ELISA. The serum concentration of SPARC was measured using commercial ELISA kit (DSP00; R&D Systems) according to manufacturer's instructions. Serum THBS2 level was measured using paired antibodies (DY1635; R&D Systems). Briefly, 96-well microplates were coated with capture antibody overnight followed by blocking with dilution buffer (1% BSA in PBS) for 2 h at room temperature. Then protein standard or serum samples (1 : 40) in dilution buffer were added to each well and incubated at room temperature for 2 h. After adding the detection antibody for 2 h incubation, the Streptavidin-HRP was added and incubated for 20 min at room temperature in the dark. The microplates were

washed four times with washing buffer (0.05% Tween-20 in PBS) after each step above. Substrate solution (1 : 1 mixture of H_2O_2 and tetramethylbenzidine) was added and incubated for 20 min at room temperature in the dark. Then 50 μl of stop solution (H_2SO_4) was added and the plates were read at 450 nm with a microplate reader (Multiskan FC, Thermo Fisher Scientific, USA).

Statistical analysis. Perseus software (Tyanova *et al*, 2016b) was used to analyse the MS data. One sample *t*-test was applied to assess the statistically significant difference between tumour and adjacent non-tumour TIF. Multiple hypothesis testing is performed using Benjamini–Hochberg correction at 5% significance level. The cutoff ratio for differentially expressed proteins was determined based on the two-fold s.d. derived from the technical repeats of pooled samples. For serum ELISA result, the SPSS software package (Version 16.0) was used for statistical analysis. Mann–Whitney *U*-test was used to compare serum candidates between HCC and healthy or benign liver disease controls. Receiver operator characteristic (ROC) curve for the combination of SPARC and THBS2 was created by the probabilities of binary logistic. The association between serum candidates and clinical characteristics was calculated based on Pearson χ^2 test or Fisher exact test. Survival curve of THBS2 was analysed with Kaplan–Meier method and *P*-values were calculated based on the log-rank test. Cox proportional hazard regression was used for multivariate analysis. All statistic tests were two-sided and statistical significance was set at $P < 0.05$.

RESULTS

Identification and quantification of proteins in TIF samples with iTRAQ experiments. Totally, 3629 proteins were identified and 3357 proteins were quantified in paired tumour and adjacent non-tumour TIFs from 16 HCC patients by iTRAQ (Figure 1A and Supplementary Table S3). Pooled tumour and adjacent non-tumour TIFs from the same 16 HCC patients were quantified in duplicate using 4-plex iTRAQ, with 2119 proteins identified and 2114 proteins quantified (Supplementary Table S4). About 98% of identified (2085) or quantified (2068) proteins in this experiment were covered by the 8-plex iTRAQ results, and the ratios$_{tumour/non-tumour}$ between pooled samples and individual samples were highly correlated ($r = 0.87$, Supplementary Figure S1). Proteins with ratios$_{tumour/non-tumour}$ greater or lower than 1.4 with Benjamin-Hochberg FDR $< 5\%$ were considered to be significantly changed (Supplementary Table S5). This resulted in 232 significantly upregulated and 257 significantly down-regulated proteins in tumour TIF compared with adjacent non-tumour TIF, which was shown in Figure 1B.

Gene ontology term enrichment analysis of TIF proteome. Gene Ontology (GO) cellular component enrichment analysis was performed with the 3629 proteins identified in TIF by DAVID. The result showed that vesicle proteins were significantly enriched in TIF (Supplementary Figure S2). The protein composition in TIF was compared with vesicle database Vesiclepedia (Kalra *et al*, 2012). It was found that about 80% of proteins identified in TIF were included in vesicle database, which indicated the important tumour microenvironment characteristics of TIF as expected. Furthermore, to identify biological function changes in HCC TIF, GO biological process enrichment analysis was performed with the differentially expressed proteins. The top ten significantly enriched GO biological processes (ranked by *P*-value) in up-regulated or downregulated proteins were shown in Figure 1C. Biological processes related to immunisation, single-organism transport, extracellular matrix organisation and blood vessel morphogenesis were significantly enriched by upregulated proteins in HCC TIF. The enriched immunity-related biological processes were
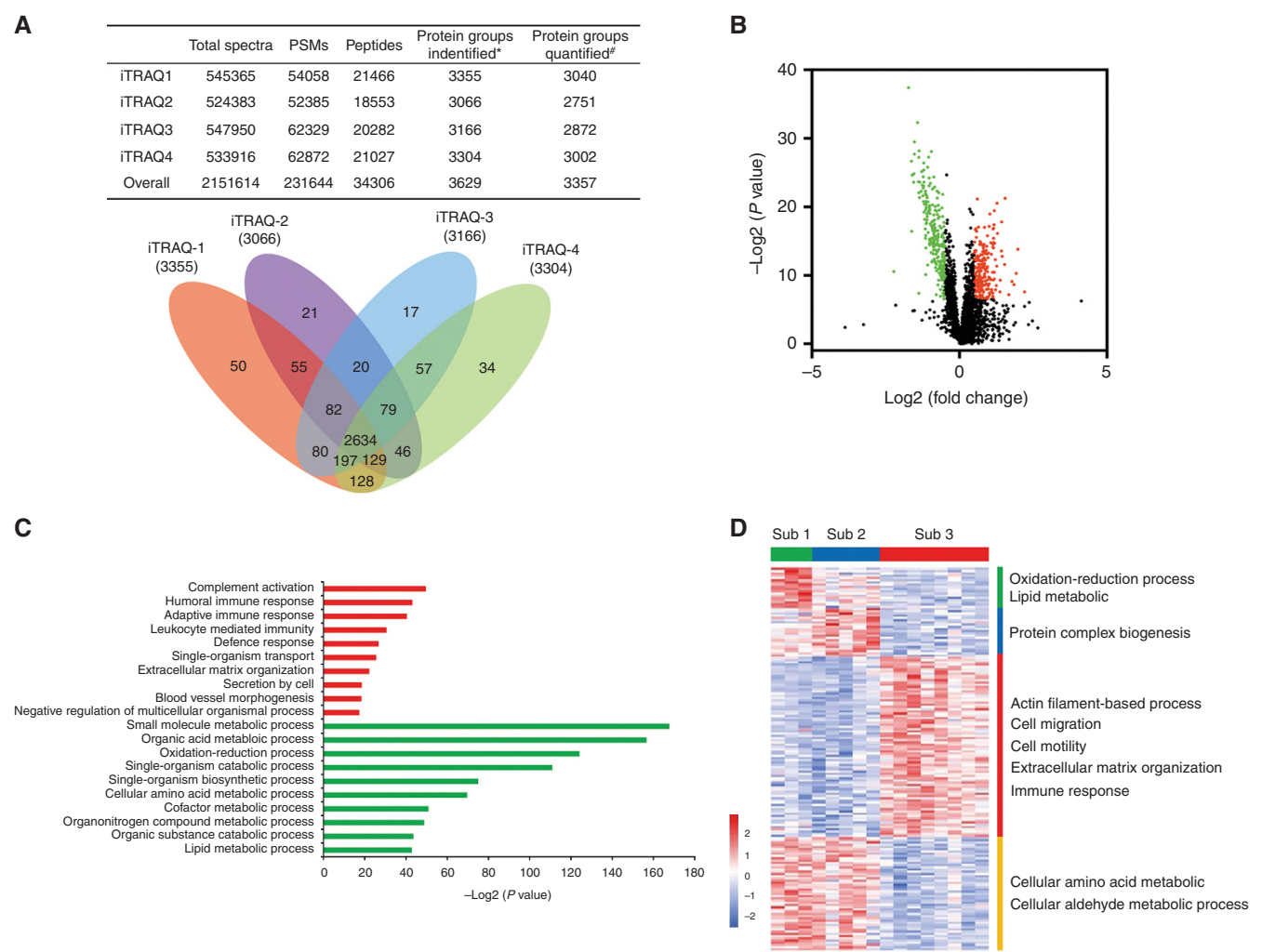
Figure 1. TIF proteome result of the iTRAQ experiment. (A) The number of spectra, peptide-spectrum matches (PSMs), distinct peptides and proteins identified in each iTRAQ batch and overall experiment. Venn diagram shows the overlap of proteins identified in 4 batches of 8-plex iTRAQ. *Proteins identified with ≥2 peptides from all samples. #Proteins identified with ≥2 peptides and quantified in at least one sample. (B) Volcano plot of differentially expressed TIF proteins. The red points represented 232 proteins that were up-regulated in HCC-TIF (adjusted P value < 0.05; fold change ≥ 1.4). The green points represented 257 proteins that were down-regulated in HCC-TIF (adjusted P value < 0.05; fold change ≤ 0.71). (C) Enriched GO biological processes in 232 upregulated (red) and 257 downregulated (green) proteins. The x axis shows the enrichment significance presented with −log2 (P-value). (D) Consensus clustering heat map of HCC TIF proteomes. K-means clustering identified three subgroups. The heat map represented the Z scores of 172 markers (based on ANOVA analysis; FDR < 0.05) that discriminate between subgroups. The biological processes enriched by subtype markers were listed on the right side.

consistent with the immune cell infiltration in tumour tissues (Supplementary Figure S3). Oxidation–reduction process and metabolic processes including small-molecule metabolic process and organic acid metabolic process were significantly enriched in the downregulated proteins.

Subtyping of HCC TIFs. Quantitative proteome data of individual HCC TIF samples were used for HCC subtyping by k-means based consensus clustering. Three subgroups were identified based on 354 proteins with no missing values and s.d. ≥ 0.7 (Supplementary Figure S4 and Supplementary Table S6). ANOVA analysis identified 172 differentially expressed proteins between any of the three subtypes (FDR < 5%). One-vs-rest t-test was used to determine subtype signatures (Supplementary Table S7). GO biological process enrichment analysis was conducted on these subtype signatures (Figure 1D and Supplementary Table S8). Oxidation reduction process and lipid metabolic process were exclusively enriched in subtype 1. Subtype 3 showed reduced metabolic related biological processes. Meanwhile, subtype 3 was

characterised by elevation of cell migration, cell motility, extracellular matrix organisation and immune response. Subtype 2 showed elevated protein complex biogenesis. In addition, the subtypes exhibited different clinical outcomes. HCC patients in subtype 3 had a significantly poor disease-free survival (DFS) compared to patients in subtype 1 ($P = 0.008$, log-rank test) and subtype 2 ($P = 0.041$, log-rank test). Although the number of samples in our study was very few, this preliminary result showed that the TIF proteome data could be used for HCC subtyping.

Confirmation of extracellular matrix-associated proteins in TIF by western blot. As the component of TIF, extracellular matrix (ECM) played important roles in cancer progression (Lu *et al*, 2012). Since extracellular matrix organisation biological process was significantly enriched in upregulated proteins in HCC TIF (Figure 1C) and in subtype 3 as elevated subtype markers (Figure 1D), we focus on ECM-associated proteins for further confirmation. Fourteen upregulated ECM proteins (Launay *et al*, 2015), which have not been reported as HCC candidate serum

biomarkers were searched in the Human Protein Atlas database (Uhlen *et al*, 2015). Three proteins including thrombospondin-2 (THBS2), Laminin subunit beta-1(LAMB1) and SPARC were found out to be upregulated in more than 50% of liver cancer tissues compared to normal liver tissues. The expression of SPARC and THBS2 in the same 16 pairs of HCC TIFs used in the iTRAQ experiment was further confirmed by western blot (Figure 2A). Samples were ordered according to each batch of the iTRAQ experiment. The result of western blot showed that both SPARC and THBS2 were significantly elevated in HCC tumour TIF compared with adjacent non-tumour TIF ($P<0.01$, Wilcoxon matched-pairs signed rank test), which was consistent with the iTRAQ result. Western blot results for the corresponding paired HCC TIFs of THBS2 were more consistent with iTRAQ ratios than those of SPARC, which might due to the more unique peptides identified by mass spectrometry for THBS2 than SPARC.

Validation of SPARC and THBS2 in the serum samples. The levels of SPARC and THBS2 were examined in the sera of 44 HCC patients and 30 healthy controls by ELISA (Supplementary Table S2). Both SPARC and THBS2 were significantly elevated in the sera of HCC compared with healthy controls ($P<0.0001$; Figure 3A). The ROC curve of SPARC for distinguishing HCC from healthy controls provided an AUC value of 0.88 (sensitivity = 80%, specificity = 93%), while the AUC of THBS2 was 0.91 (sensitivity = 80%, specificity = 93%). When these two candidates were combinedly used (the probabilities of binary logistic were calculated by SPSS and used to create ROC curve), higher accuracy in distinguishing HCC from healthy controls (AUC = 0.97, sensitivity = 86%, specificity = 100%) was achieved. For AFP negative ($<20\,\mathrm{ng\,ml^{-1}}$) samples ($n=22$), serum concentration of SPARC and THBS2 in HCC were also significantly higher than that in healthy controls ($P<0.0001$, Figure 3B). The ROC curve showed that SPARC (AUC = 0.88, sensitivity = 77%, specificity = 93%) and THBS2 (AUC = 0.90, sensitivity = 86%, specificity = 87%) could also distinguish AFP-negative HCC patients from healthy controls. Combination of serum SPARC and THBS2 achieved higher accuracy for distinguishing AFP-negative HCC patients from healthy controls (AUC = 0.95, sensitivity = 91%, specificity = 93%; Figure 3C).

Besides this, the levels of SPARC and THBS2 were also examined in the sera of 17 benign liver disease patients (Supplementary Figure S5). It was found that the levels of SPARC in the sera of HCC were significantly lower than those in benign liver disease patients ($P<0.0001$), while THBS2 in the sera of HCC were significantly higher than those in benign liver disease patients ($P=0.0004$). Combination of serum SPARC and THBS2 could distinguish HCC patients from benign liver disease patients (AUC = 0.93, sensitivity = 80%, specificity = 94%). Results above indicated that the combination of serum SPARC and THBS2 could be used for HCC screening in asymptomatic healthy controls and benign liver disease patients.
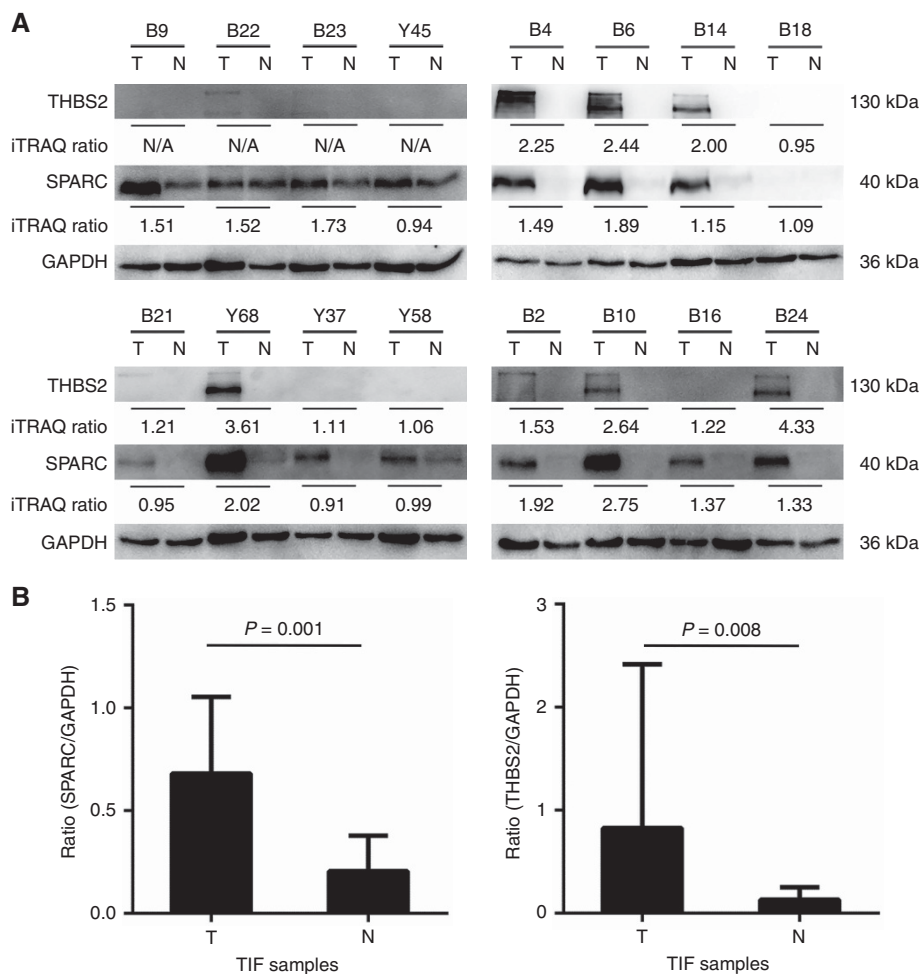

Figure 2. Western blot confirmation of candidate proteins in TIF samples. (**A**) Western blot of SPARC and THBS2 in 16 paired HCC TIF samples. Two THBS2 bands at 130 kDa and 170 kDa were detected in TIF samples. (**B**) The levels of SPARC and THBS2 were significantly elevated in HCC TIF compared with adjacent non-tumour TIF ($P=0.001$ or $P=0.008$, Wilcoxon matched-pairs signed rank test).
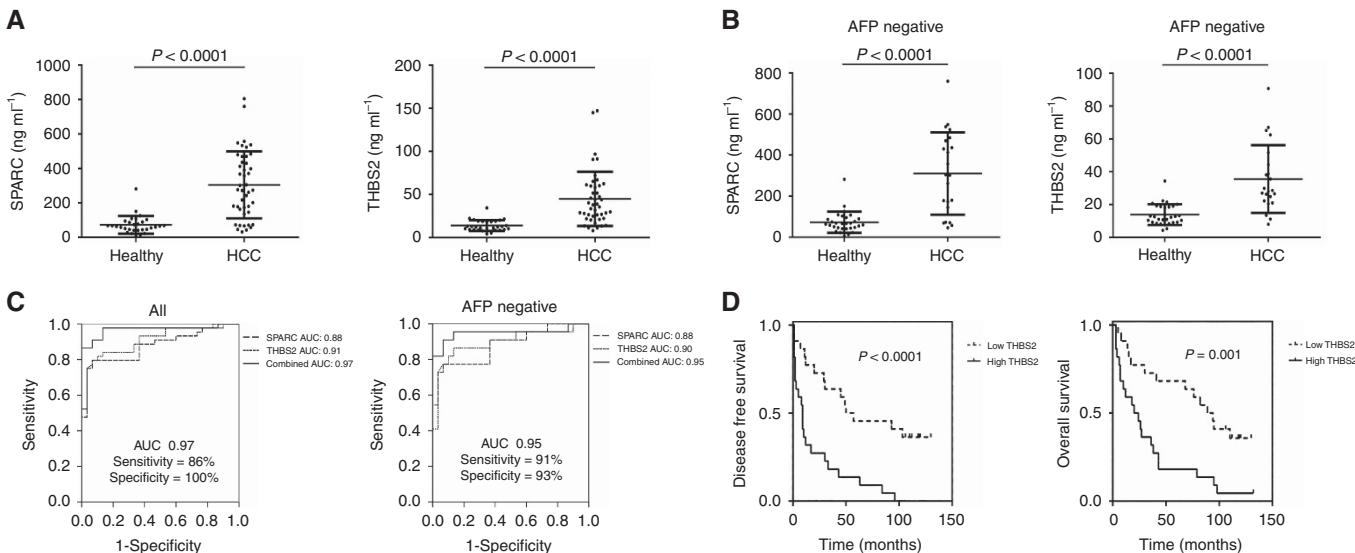
Figure 3. ELISA validation result of serum SPARC and THBS2. (A) Serum levels of SPARC (left) and THBS2 (right) were significantly elevated in HCC patients ($n = 44$) compared with healthy controls ($n = 30$; $P < 0.0001$, Mann–Whitney U-test). The horizontal lines indicate means ± s.d. (B) SPARC (left) and THBS2 (right) concentration in healthy controls ($n = 30$) and AFP-negative HCC patients ($n = 22$). HCC patients had higher serum SPARC and THBS2 levels than healthy controls ($P < 0.0001$, Mann–Whitney U-test). (C) ROC curves for SPARC and THBS2 to discriminate HCC (left) or AFP-negative HCC (right) patients with healthy controls. The combination of serum SPARC and THBS2 achieve better performance than SPARC or THBS2 alone in discriminating HCC or AFP-negative HCC patients from healthy controls. (D) Kaplan-Meier analysis of disease-free survival (left) and overall survival (right) of HCC patients with high or low serum THBS2 levels (cutoff 36.9 ng ml^{-1} according to the median value). Higher serum THBS2 was significantly correlated with short DFS and OS in HCC patients (log-rank test).

Table 2. Association between serum candidates and clinical characteristics in HCC patients

Characteristics	Cases	SPARC			THBS2		
		Low[a]	High[a]	P-value[b]	Low[a]	High[a]	P-value[b]
Age (years)							
<60	34	15	19	0.150	16	18	0.472
⩾60	10	7	3		6	4	
Gender							
Male	37	19	18	1.000	19	18	1.000
Female	7	3	4		3	4	
Tumour size (cm)							
<5	14	8	6	0.517	8	6	0.517
⩾5	30	14	16		14	16	
Microvascular invasion							
Positive	9	5	4	1.000	1	8	**0.021**
Negative	35	17	18		21	14	
TNM stage							
I + II	32	14	18	0.176	19	13	**0.042**
III + IV	12	8	4		3	9	
AFP (ng ml^{-1})							
<20	22	10	12	0.546	14	8	0.070
⩾20	22	12	10		8	14	

NOTE: statistically significant values are in bold type.
[a]For SPARC and THBS2, median values were used as the cutoff point to demarcate high- and low-expression groups.
[b]P-value was calculated based on Pearson χ^2-test or Fisher exact test and were statistically significant at <0.05.

Prognostic value of THBS2 in HCC patients. The correlations between serum concentrations of these two candidates and clinical characteristics of HCC patients were examined. Serum levels of SPARC and THBS2 were divided into high and low groups based on their median values in 44 HCC patients. The result showed that high level of serum THBS2 was significantly correlated with microvascular invasion ($P = 0.021$) and TNM stage ($P = 0.042$) in

HCC patients (Table 2). More importantly, the multivariate Cox regression analysis of HCC patients showed that serum THBS2 level ($P = 0.016$), together with microvascular invasion, were independent risk factors for DFS in HCC patients (Table 3). Then the association between serum THBS2 and HCC patients' prognosis was assessed. Kaplan–Meier survival analysis revealed that HCC patients with high THBS2 levels had significantly shorter

Table 3. Multivariate Cox regression analysis of serum THBS2 expression and HCC patients' survival

Variable	Disease-free survival			Overall survival		
	HR	95% CI	P value	HR	95% CI	P value
Age (years)						
⩾60 vs <60	1.684	0.740–3.829	0.214	2.395	1.015–5.647	**0.046**
Gender						
Male vs female	0.517	0.183–1.466	0.215	0.556	0.183–1.690	0.301
Tumour size (cm)						
⩾5 vs <5	0.934	0.406–2.149	0.873	1.048	0.457–2.405	0.912
Microvascular invasion						
Positive vs Negative	4.015	1.372–11.746	**0.011**	7.896	2.235–27.886	**0.001**
TNM stage						
III+IV vs I+II	2.168	0.883–5.324	0.092	3.134	1.198–8.202	**0.020**
AFP (ng ml^{-1})						
⩾20 vs <20	1.450	0.589–3.567	0.419	1.783	0.734–4.334	0.202
THBS2 (ng ml^{-1})[a]						
⩾36.9 vs <36.9	2.690	1.203–6.012	**0.016**	2.070	0.896–4.779	0.088

Abbreviations: AFP = alpha-fetoprotein; CI = confidence interval; HR = hazard ratio. Statistically significant values are in bold type.
[a]For THBS2, median values were used as the cut-off point to demarcate high and low expression groups.

DFS ($P<0.0001$) and overall survival (OS) ($P=0.001$) than those with low THBS2 levels (Figure 3D). These results indicated that THBS2 could be used as a candidate serum biomarker for HCC prognosis.

DISCUSSION

As the body fluid exists between tumour cells and capillary blood, TIF contains proteins secreted or shed from tumour cells and surrounding stromal cells, which could reflect the pathological alterations within the tumour microenvironment. Our previous works proved that liver TIF does not have obvious high-abundance proteins, which makes it easier to identify tumour-derived low-abundance proteins. Although the concentration of proteins in TIF will be greatly diluted in blood, which may lead to failed validation of candidates in blood, TIF is still a good source for biomarker discovery. In this study, we profiled the proteomes of 16 paired HCC TIF samples and identified 3629 proteins, which is the largest proteomic dataset of HCC TIF with high confidence to our knowledge.

It is known that sometimes the alteration of protein levels in serum may not be consistent with that in tumour tissue. For example, Na's group (Na et al, 2009, 2013) reported that the level of liver carboxylesterase 1 (CES1) was elevated in the plasma of HCC patients, but was remarkably downregulated in HCC tumour tissues. As CES1 was secreted by tumour cells, Na's group proposes a hypothesis that the opposite alteration of CES1 in tissue and plasma of HCC patients might due to the enhanced secretion in tumour tissue compared with control tissue. In our study, the level of CES1 was elevated in 62.5% (10/16) of HCC TIF samples, which indicated that the level change of secreted proteins in TIF was more consistent with that in serum sometimes. In this point, TIF is an ideal source for biomarker discovery.

ECM proteins, which play roles in cell adhesion and cell-to-cell communication, have been proved to be potential cancer biomarkers (Paradis et al, 2005; Shang et al, 2012; Lucarelli et al, 2013; Naba et al, 2014). In our study, ECM proteins were found upregulated and ECM organisation biological process was significantly enriched in HCC TIF and in subtype 3 HCC patients. Two ECM proteins, SPARC and THBS2 were selected for further

validation. SPARC is a glycoprotein involved in cell growth regulation through interactions with the ECM and cytokines. The role of SPARC in cancer is complex because SPARC could promote tumour progression in some certain cancers such as melanoma and glioma, while may function as a tumour suppressor in other types of cancer such as ovarian cancer and colorectal cancer (Tai and Tang, 2008). The role of SPARC in HCC remains unclear. Lau et al (2006) found that overexpression of SPARC in mouse xenograft models resulted in the delayed growth of HCC. Atorrasagasti et al (2010) showed that over-expression of SPARC reduced the tumorigenicity of HCC cells. However, Deng et al (2016) reported that as a target of miR-211, uncontrolled SPARC over-expression might be related with HCC progression. Although the role of SPARC in HCC is not clear, SPARC was reported to be upregulated in HCC tissues consistently (Le Bail et al, 1999; Lau et al, 2006). Our results for the first time demonstrated that SPARC was elevated both in HCC TIFs and sera. And serum SPARC could efficiently distinguish HCC or AFP-negative HCC patients from healthy controls.

Thrombospondins are a family of secreted glycoproteins mediating antiangiogenesis. In our data, thrombospondin-1 (THBS1) and thrombospondin-2 (THBS2) were both up-regulated in HCC TIF. THBS1 has been reported as a potential diagnostic and prognostic marker for HCC (Poon et al, 2004; Chen et al, 2011) and high expression of THBS1 might be associated with HCC invasiveness (Poon et al, 2004). However, there is no study about THBS2 in HCC yet. The overexpression of THBS2 in tumour tissue was correlated with poor prognosis in renal cancer (Boguslawska et al, 2016), ovarian cancer (Cheon et al, 2014) and oral cavity squamous cell carcinoma (Hsu et al, 2014). However, upregulation of THBS2 in gastric cancer (Sun et al, 2014) tissue was correlated with favourable prognosis, suggesting that THBS2 might play different roles in different kinds of cancer. Our study discovered the upregulation of THBS2 in HCC TIF. Western blot result confirmed the elevation of THBS2 in HCC TIF and showed that about half of HCC TIFs had higher level of THBS2. The individual differences of THBS2 expression might due to the different subtype of HCC patients as THBS2 was a marker of subtype 3 HCC. ELISA result showed that the serum level of THBS2 was elevated in HCC patients. THBS2 could distinguish HCC or AFP-negative HCC from healthy controls. We also found that the serum level of THBS2 in HCC patients was correlated with microvascular invasion and advanced TNM stage. Survival analysis showed that high serum THBS2 level was associated with poor prognosis and THBS2 was an independent risk factor for HCC recurrence. To our knowledge, this is the first time that serum THBS2 was found to be a candidate biomarker for HCC. Besides, the differential expression of serum SPARC and THBS2 between HCC and benign liver disease patients were also detected. Further analysis found that combined serum level of SPARC and THBS2 could distinguish HCC from healthy or benign liver disease group, suggesting the potential value of these two candidate HCC serum biomarkers for further large-scale clinical validation.

In conclusion, we profiled 16 pairs of HCC TIFs with iTRAQ technique. Two ECM proteins, SPARC and THBS2 were found to be candidate serum biomarkers for HCC screening. And serum THBS2 was also found to be a novel indicator for poor prognosis of HCC. This study showed that TIF could reflect the pathological changes of HCC tumour microenvironment and serve for serum biomarker discovery with the in-depth proteomic analysis.

ACKNOWLEDGEMENTS

Technology Program (Z141100000214014, Z161100002616036) and State Key Laboratory of Proteomics (SKLP-0201410).

REFERENCES

(2010) Clinical Practice Guidelines for Hepatocellular Carcinoma - The Japan Society of Hepatology 2009 update. *Hepatol Res* **40**(Suppl 1): 2–144.

Atorrasagasti C, Malvicini M, Aquino JB, Alaniz L, Garcia M, Bolontrade M, Rizzo M, Podhajcer OL, Mazzolini G (2010) Overexpression of SPARC obliterates the in vivo tumorigenicity of human hepatocellular carcinoma cells. *Int J Cancer* **126**: 2726–2740.

Boguslawska J, Kedzierska H, Poplawski P, Rybicka B, Tanski Z, Piekielko-Witkowska A (2016) Expression of genes involved in cellular adhesion and extracellular matrix remodeling correlates with poor survival of patients with renal cancer. *J Urol* **195**: 1892–1902.

Celis JE, Gromov P, Cabezon T, Moreira JM, Ambartsumian N, Sandelin K, Rank F, Gromova I (2004) Proteomic characterization of the interstitial fluid perfusing the breast tumor microenvironment: a novel resource for biomarker and therapeutic target discovery. *Mol Cell Proteomics* **3**: 327–344.

Chen R, Tan Y, Wang M, Wang F, Yao Z, Dong L, Ye M, Wang H, Zou H (2011) Development of glycoprotein capture-based label-free method for the high-throughput screening of differential glycoproteins in hepatocellular carcinoma. *Mol Cell Proteomics* **10**: M110 006445.

Cheon DJ, Tong Y, Sim MS, Dering J, Berel D, Cui X, Lester J, Beach JA, Tighiouart M, Walts AE, Karlan BY, Orsulic S (2014) A collagen-remodeling gene signature regulated by TGF-beta signaling is associated with metastasis and poor survival in serous ovarian cancer. *Clin Cancer Res* **20**: 711–723.

Daniele B, Bencivenga A, Megna AS, Tinessa V (2004) Alpha-fetoprotein and ultrasonography screening for hepatocellular carcinoma. *Gastroenterology* **127**: S108–S112.

Deng B, Qu L, Li J, Fang J, Yang S, Cao Z, Mei Z, Sun X (2016) MiRNA-211 suppresses cell proliferation, migration and invasion by targeting SPARC in human hepatocellular carcinoma. *Sci Rep* **6**: 26679.

Hsu CW, Yu JS, Peng PH, Liu SC, Chang YS, Chang KP, Wu CC (2014) Secretome profiling of primary cells reveals that THBS2 is a salivary biomarker of oral cavity squamous cell carcinoma. *J Proteome Res* **13**: 4796–4807.

Huang da W, Sherman BT, Lempicki RA (2009) Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**: 44–57.

Kalra H, Simpson RJ, Ji H, Aikawa E, Altevogt P, Askenase P, Bond VC, Borras FE, Breakefield X, Budnik V, Buzas E, Camussi G, Clayton A, Cocucci E, Falcon-Perez JM, Gabrielsson S, Gho YS, Gupta D, Harsha HC, Hendrix A, Hill AF, Inal JM, Jenster G, Kramer-Albers EM, LIM SK, Llorente A, Lotvall J, Marcilla A, Mincheva-Nilsson L, Nazarenko I, Nieuwland R, Nolte-'T Hoen EN, Pandey A, Patel T, Piper MG, Pluchino S, Prasad TS, Rajendran L, Raposo G, Record M, Reid GE, Sanchez-Madrid F, Schiffelers RM, Siljander P, Stensballe A, Stoorvogel W, Taylor D, Thery C, Valadi H, Van Balkom BW, Vazquez J, Vidal M, Wauben MH, Yanez-Mo M, Zoeller M, Mathivanan S (2012) Vesiclepedia: a compendium for extracellular vesicles with continuous community annotation. *PLoS Biol* **10**: e1001450.

Kikuchi T, Hassanein M, Amann JM, Liu Q, Slebos RJ, Rahman SM, Kaufman JM, Zhang X, Hoeksema MD, Harris BK, Li M, Shyr Y, Gonzalez AL, Zimmerman LJ, Liebler DC, Massion PP, Carbone DP (2012) In-depth proteomic analysis of nonsmall cell lung cancer to discover molecular targets and candidate biomarkers. *Mol Cell Proteomics* **11**: 916–932.

Lau CP, Poon RT, Cheung ST, Yu WC, Fan ST (2006) SPARC and Hevin expression correlate with tumour angiogenesis in hepatocellular carcinoma. *J Pathol* **210**: 459–468.

Launay G, Salza R, Multedo D, Thierry-Mieg N, Ricard-Blum S (2015) MatrixDB, the extracellular matrix interaction database: updated content,

a new navigator and expanded functionalities. *Nucleic Acids Res* **43**: D321–D327.

Le Bail B, Faouzi S, Boussarie L, Guirouilh J, Blanc JF, Carles J, Bioulac-Sage P, Balabaud C, Rosenbaum J (1999) Osteonectin/SPARC is overexpressed in human hepatocellular carcinoma. *J Pathol* **189**: 46–52.

Lu P, Weaver VM, Werb Z (2012) The extracellular matrix: a dynamic niche in cancer progression. *J Cell Biol* **196**: 395–406.

Lucarelli G, Rutigliano M, Bettocchi C, Palazzo S, Vavallo A, Galleggiante V, Trabucco S, Di Clemente D, Selvaggi FP, Battaglia M, Ditonno P (2013) Spondin-2, a secreted extracellular matrix protein, is a novel diagnostic biomarker for prostate cancer. *J Urol* **190**: 2271–2277.

Mertins P, Mani DR, Ruggles KV, Gillette MA, Clauser KR, Wang P, Wang X, Qiao JW, Cao S, Petralia F, Kawaler E, Mundt F, Krug K, Tu Z, Lei JT, Gatza ML, Wilkerson M, Perou CM, Yellapantula V, Huang KL, Lin C, Mclellan MD, Yan P, Davies SR, Townsend RR, Skates SJ, Wang JJ, Zhang B, Kinsinger CR, Mesri M, Rodriguez H, Ding L, Paulovich AG, Fenyo D, Ellis MJ, Carr SA, Nci C (2016) Proteogenomics connects somatic mutations to signalling in breast cancer. *Nature* **534**: 55–62.

Mittal S, El-Serag HB (2013) Epidemiology of hepatocellular carcinoma: consider the population. *J Clin Gastroenterol* **47**(Suppl): S2–S6.

Na K, Jeong SK, Lee MJ, Cho SY, Kim SA, Lee MJ, Song SY, Kim H, Kim KS, Lee HW, Paik YK (2013) Human liver carboxylesterase 1 outperforms alpha-fetoprotein as biomarker to discriminate hepatocellular carcinoma from other liver diseases in Korean patients. *Int J Cancer* **133**: 408–415.

Na K, Lee EY, Lee HJ, Kim KY, Lee H, Jeong SK, Jeong AS, Cho SY, Kim SA, Song SY, Kim KS, Cho SW, Kim H, Paik YK (2009) Human plasma carboxylesterase 1, a novel serologic biomarker candidate for hepatocellular carcinoma. *Proteomics* **9**: 3989–3999.

Naba A, Clauser KR, Lamar JM, Carr SA, Hynes RO (2014) Extracellular matrix signatures of human mammary carcinoma identify novel metastasis promoters. *Elife* **3**: e01308.

Omata M, Lesmana LA, Tateishi R, Chen PJ, Lin SM, Yoshida H, Kudo M, Lee JM, Choi BI, Poon RT, Shiina S, Cheng AL, Jia JD, Obi S, Han KH, Jafri W, Chow P, Lim SG, Chawla YK, Budihusodo U, Gani RA, Lesmana CR, Putranto TA, Liaw YF, Sarin SK (2010) Asian Pacific Association for the Study of the Liver consensus recommendations on hepatocellular carcinoma. *Hepatol Int* **4**: 439–474.

Paradis V, Degos F, Dargere D, Pham N, Belghiti J, Degott C, Janeau JL, Bezeaud A, Delforge D, Cubizolles M, Laurendeau I, Bedossa P (2005) Identification of a new marker of hepatocellular carcinoma by serum protein profiling of patients with chronic liver diseases. *Hepatology* **41**: 40–47.

Poon RT, Chung KK, Cheung ST, Lau CP, Tong SW, Leung KL, Yu WC, Tuszynski GP, Fan ST (2004) Clinical significance of thrombo-spondin 1 expression in hepatocellular carcinoma. *Clin Cancer Res* **10**: 4150–4157.

Shang S, Plymoth A, Ge S, Feng Z, Rosen HR, Sangrajrang S, Hainaut P, Marrero JA, Beretta L (2012) Identification of osteopontin as a novel marker for early hepatocellular carcinoma. *Hepatology* **55**: 483–490.

Sherman M (2005) Hepatocellular carcinoma: epidemiology, risk factors, and screening. *Semin Liver Dis* **25**: 143–154.

Sun R, Wu J, Chen Y, Lu M, Zhang S, Lu D, Li Y (2014) Down regulation of Thrombospondin2 predicts poor prognosis in patients with gastric cancer. *Mol Cancer* **13**: 225.

Sun W, Ma J, Wu S, Yang D, Yan Y, Liu K, Wang J, Sun L, Chen N, Wei H, Zhu Y, Xing B, Zhao X, Qian X, Jiang Y, He F (2010) Characterization of the liver tissue interstitial fluid (TIF) proteome indicates potential for application in liver disease biomarker discovery. *J Proteome Res* **9**: 1020–1031.

Tai IT, Tang MJ (2008) SPARC in cancer biology: its role in cancer progression and potential for therapy. *Drug Resist Updat* **11**: 231–246.

Teng PN, Hood BL, Sun M, Dhir R, Conrads TP (2011) Differential proteomic analysis of renal cell carcinoma tissue interstitial fluid. *J Proteome Res* **10**: 1333–1342.

Tyanova S, Albrechtsen R, Kronqvist P, Cox J, Mann M, Geiger T (2016a) Proteomic maps of breast cancer subtypes. *Nat Commun* **7**: 10259.

Tyanova S, Temu T, Sinitcyn P, Carlson A, Hein MY, Geiger T, Mann M, Cox J (2016b) The Perseus computational platform for comprehensive analysis of (prote)omics data. *Nat Methods* **13**: 731–740.

Uhlen M, Fagerberg L, Hallstrom BM, Lindskog C, Oksvold P, Mardinoglu A, Sivertsson A, Kampf C, Sjostedt E, Asplund A, Olsson I, Edlund K, Lundberg E, Navani S, Szigyarto CA, Odeberg J, Djureinovic D, Takanen JO, Hober S, Alm T, Edqvist PH, Berling H, Tegel H, Mulder J, Rockberg J, Nilsson P, Schwenk JM, Hamsten M, Von Feilitzen K, Forsberg M, Persson L, Johansson F, Zwahlen M, Von Heijne G, Nielsen J, Ponten F (2015) Proteomics. Tissue-based map of the human proteome. *Science* **347**: 1260419.

Weng TY, Wang CY, Hung YH, Chen WC, Chen YL, Lai MD (2016) Differential expression pattern of THBS1 and THBS2 in lung cancer: clinical outcome and a systematic-analysis of microarray databases. *PLoS One* **11**: e0161007.

Wong SL, Sukkar MB (2017) The SPARC protein: an overview of its role in lung cancer and pulmonary fibrosis and its potential role in chronic airways disease. *Br J Pharmacol* **174**: 3–14.

Xie Y, Chen L, Lv X, Hou G, Wang Y, Jiang C, Zhu H, Xu N, Wu L, Lou X, Liu S (2016) The levels of serine proteases in colon tissue interstitial fluid and serum serve as an indicator of colorectal cancer progression. *Oncotarget* **7**: 32592–32606.

Supplementary Information accompanies this paper on British Journal of Cancer website (http://www.nature.com/bjc)