

## Identification and validation of circulating biomarkers for detection of liver cancer with antibody array

Shuangzhe ZHANG<sup>1,2,3,\*</sup>, Chunhui GAO<sup>4,†</sup>, Qi ZHOU<sup>5,6,†</sup>, Lipai CHEN<sup>4,†</sup>, Wei HUANG<sup>1</sup>, Gordon Fan HUANG<sup>7</sup>, Hao TANG<sup>1,7</sup>, Xuedong SONG<sup>1</sup>, Zhuo ZHANG<sup>1</sup>, Kelly WHITTAKER<sup>7</sup>, Xiaofeng CHEN<sup>3,8,9</sup>, Ruo-Pan HUANG<sup>1,2,7,\*</sup>

<sup>1</sup>Raybiotech Co., Ltd., Guangzhou, Guangdong, China; <sup>2</sup>South China Biochip Research Center, Guangzhou, Guangdong, China; <sup>3</sup>Department of Biomedical Engineering, School of Materials Science and Engineering, South China University of Technology, Guangzhou, Guangdong, China; <sup>4</sup>Affiliated Cancer Hospital and Institute of Guangzhou Medical University, Guangzhou, Guangdong, China; <sup>5</sup>Department of Liver Surgery, The First Affiliated Hospital of Sun Yat-Sen University, Guangzhou, Guangdong, China; <sup>6</sup>Department of General Surgery, Hui Ya Hospital of The First Affiliated Hospital Sun Yat-Sen University, Huizhou, Guangdong, China; <sup>7</sup>RayBiotech Life Inc., Peachtree Corners, Georgia, United States; <sup>8</sup>National Engineering Research Center for Tissue Restoration and Reconstruction, Guangzhou Higher Education Mega Centre, Guangzhou, Guangdong, China; <sup>9</sup>Key Laboratory of Biomedical Engineering of Guangdong Province, South China University of Technology, Guangzhou Higher Education Mega Centre, Guangzhou, Guangdong, China

\*Correspondence: rhuang@raybiotech.com

†Contributed equally to this work.

Received June 6, 2022 / Accepted December 9, 2022

The aim of this study was to find new protein biomarkers that could be used to detect hepatocellular carcinoma (HCC) in the serum. We identified 11 proteins in the tissue that could be used to classify samples from HCC and control subjects. The 11 identified tissue biomarkers were combined with 10 commonly used serum HCC biomarkers for further verification in a large number of serum samples from HCC patients and healthy controls. 17 of the 21 prospective serum biomarkers were determined to be differentially expressed through collinearity and significance analysis. Through the method of supervised learning, a random forest model was constructed to reduce the dimensionality of the number of differentially expressed proteins, and finally, 4 differentially expressed proteins were identified: AFP, GDF15, CEACAM-1, and MMP-9, and suggested to have potential application in clinical diagnosis of HCC.

*Key words: liver cancer; protein biomarkers; antibody array; random forest; tissue and serum detection; combination of biomarkers*

Hepatocellular carcinoma (HCC) is the 5<sup>th</sup> most common cancer, but the 3<sup>rd</sup> leading cause of cancer death globally with approximately 700,000 fatalities annually [1, 2]. HCC ranks as the 2<sup>nd</sup> most malignant tumor in China, accounting for about 55% of global HCC cases. The high mortality of HCC is always closely associated with late detection. The onset of HCC is often not accompanied by obvious clinical symptoms, and it is often difficult for patients to detect it at an early stage [3]. Most patients with HCC are already in the middle and advanced stages when they are diagnosed.

The current clinical main methods for the detection of HCC are magnetic resonance imaging (MRI) [4], ultrasound (US) [5], and computed tomography (CT) [6] scans to detect lesions [7]. In contrast with these detection methods, biomarkers from peripheral blood are a widely accepted new way to detect primary HCC and its metastases. The liver secretes many proteins into the blood, allowing for a

non-invasive collection of proteins for analysis. Peripheral blood biomarker detection has the additional advantages of convenience and low cost compared to the commonly used imaging methods. Various proteomic methodologies have been proposed to identify proteins that are altered in the serum of those with HCC. Using such methods, proteins such as peroxiredoxin 3, osteopontin, and alpha-fetoprotein (AFP) have been identified as potential markers of HCC [8], with upregulation of these proteins observed in HCC patient samples compared to healthy individuals or those with liver disease [9]. AFP lacks the specificity and sensitivity to stand alone as a powerful biomarker. However, AFP combined with US screening significantly increased the sensitivity of US screens [10] and now becomes one of the most widely used screening methods for HCC. Another serum protein that has shown potential as a biomarker is des-gamma carboxyprothrombin (DCP) [11], and studies suggest it to

be a more powerful biomarker than AFP for larger tumors as well as those arising from viral etiology. Recent experiments have begun to utilize combinations of protein markers to create more sensitive biomarker panels, for example combining AFP with another serum protein, fibronectin 1 [12]. This multi-marker panel approach illustrates that detection performance can be improved by integrating separately characterized protein biomarkers. 60–70% of patients with primary liver cancer have elevated AFP levels, as well as testicular cancer, ovarian tumors, pancreatic cancer, gastric cancer, bowel cancer, and lung cancer. Therefore, it is critical to developing novel protein biomarkers to match the application of AFP. It can be divided into two cases, one is that when AFP is not expressed, the novel biomarkers can detect the occurrence and development of liver cancer, and the other is that when AFP is present, the novel biomarkers are needed to rule out other cancers mentioned above. In order to identify novel and effective biomarkers, we first screened out differential proteins from tissue biomarkers. Since the differential protein in the tissue does not mean that it will enter the peripheral blood and can be detected, therefore, we screened the differential protein in the tissue and then verified it in the serum to achieve real-time non-invasive detection [13, 14].

At present, partial surgical resection and liver transplantation are the main methods of radical treatment for HCC, but the recurrence rate still reaches 50–70% after 5 years. The 5-year survival rate for primary HCC patients diagnosed at advanced stages is less than 5%, while for those diagnosed at early stages, the survival rate can reach 70–80%, which highlights the importance of early detection to improve the overall survival. Besides, cancer metastasis is now recognized as a process of interaction between tumor cells and the host microenvironment, which can affect tumor biology and promote or inhibit tumor cell colonization in target organs. With the emergence of high-throughput research methods such as microarray technology, a large number of molecules, gene expression profiles, and signaling pathways derived from the microenvironment and involved in the specific metastasis of tumor organs have been discovered. Early detection and identification are important for monitoring and preventing the occurrence of tumor metastasis. Currently, there is a lack of an effective biomarker or biomarker panel for early detection of HCC, metastasis, and prognosis. Therefore, finding new biomarkers for HCC will be of great value [15].

## Patients and methods

**Patient samples of tissues and sera.** The patients' samples were collected at the Department of Surgery, The First Affiliated Hospital, Sun Yat-Sen University as described in the previous publication [16]. HCC tissue and para-cancerous tissue were obtained from 25 patients (Table 1). In addition, serum samples from 76 HCC patients and 120 controls (80 healthy, 10 esophageal cancer, 10 lung cancer, and 20 gastric cancer) were obtained for validation (Table 2). The detection

of HCC was determined pathologically and immunohistochemically according to the WHO classification. Prior to the collection of patient information and serum samples for analysis, written informed consent was obtained from each participant, and this study was approved by the institutional ethics committee of Sun Yat-Sen University.

**Screening of candidate biomarkers.** A total of 25 HCC tissues and 25 para-cancerous tissues of 25 patients were subjected to RayBiotech antibody array AAH-CYT-G4000 with 274 secreted proteins for the candidate biomarker screening. This array is printed on a glass slide and is essentially a semi-quantitative, high-throughput sandwich-based ELISA. Each capture antibody was printed in duplicate on the glass for detection. All sample handling and microarray assays were performed in accordance with the operating instructions.

Tissue samples were cut into 1–3 mm<sup>3</sup> pieces, 500 µl of tissue lysis buffer was added and a homogenizer was used for tissue lysis. 100 µl blocking buffer was added into each well and incubated at room temperature for 30 min to block slides. Arrays were incubated with samples at room temperature for 2 h. After washing, the arrays were incubated with a biotinylated antibody cocktail for 2 h at room temperature. After washing, Cy3 equivalent dye-streptavidin was added, and the fluorescent signal was visualized using a laser scanner equipped with a Cy3 wavelength (green channel) (InnoScan 300 Microarray Scanner, Innopsys, France). The data were normalized using the RayBiotech analysis tool, an array-specific, Excel-based program that performs sophisticated data analysis on the raw numerical data extracted from the array scan.

Serum samples from 76 HCC patients and 120 controls were diluted 2-fold and loaded on the array directly. A custom antibody array was developed for the detection of 21 proteins (11 proteins with significant differences screened from HCC tissues and 10 common HCC serum biomarkers reported in the literature). The subsequent experimental steps are the same as the antibody array utilized for the tissue sample indicated above.

**Model construction.** Four different models (LR, LDA, RF, and SVM) were trained with R package 'caret'. The data set (76 HCC samples, 120 controls) was split into two subsets at a ratio of 4:1. The larger (4/5) subset was used for model training, while the smaller one (1/5) served as test/validating samples. The 4 models were trained using a scheme of 4-fold cross-validation with 5 repeats. The performance of these models was evaluated by accuracy and kappa values from cross-validation during model training, and the AUC analysis on model prediction in the test dataset.

**Statistical analysis.** The concentration of all tested samples was calculated using a linear regression curve fitted to the standard data points. A t-test in SPSS 20.0 was used to analyze the protein level differences between HCC patients and healthy volunteers. The sensitivity and specificity of all biomarkers for HCC detection were evaluated by receiver

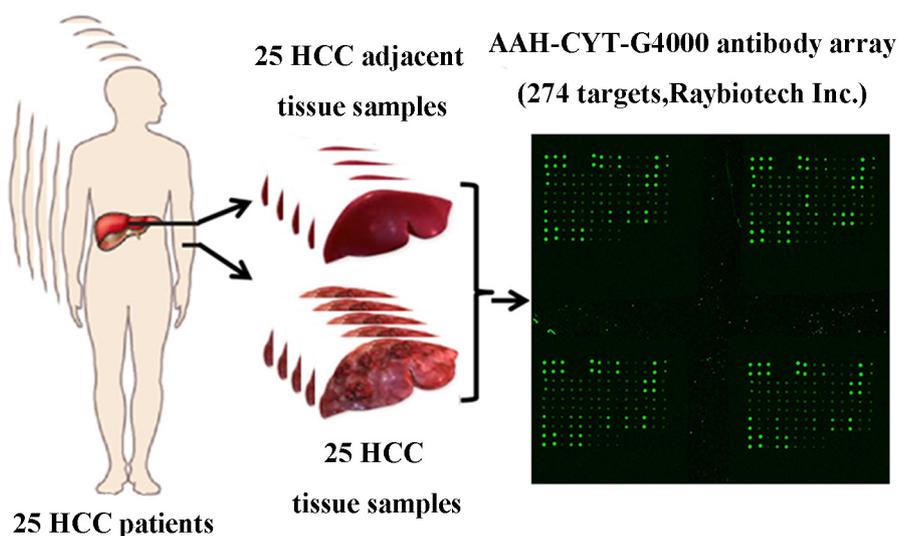
operating characteristic (ROC) curves and areas under the curves (AUC) with a 95% confidence interval (CI). The comparison ROC was conducted by GraphPad (version 8.0). The best cut-off value for detection was determined by maximizing the specificity and sensitivity at a 95% CI. A two-tailed p-value less than 0.05 was considered significant. All figures were generated using GraphPad Prism version 8.0 for Windows.

## Results

**The antibody array identified 11 candidate biomarkers in HCC.** An antibody array AAH-CYT-G4000 was used to detect the expression levels of 274 proteins in the cancer tissues and para-cancer tissues of a cohort of 25 HCC patients (Figure 1). SAM (Significant Analysis of Microarray, FDR <5%) resulted in 69 differential proteins between HCC tissues and para-cancerous tissues, of which 15 were upregulated and 54 were downregulated (Supplementary Table S1). A Wilcoxon test ( $p < 0.05$ ) found 76 proteins were significantly differentially expressed in HCC tissues and para-cancer tissues, of which 12 were upregulated and 64 were downregulated (Supplementary Table S2). A paired t-test ( $p < 0.05$ ) yielded 67 differentially expressed proteins, of which 13 were upregulated and 54 were downregulated (Supplementary Table S3). Adding the condition  $|d\text{-score}| > 3$  in SAM and intersecting with the results of the Wilcoxon and paired t-test, 33 differentially expressed proteins were identified, of which 4 were upregulated and 29 were downregulated (Supplementary Table S4, Figure 2A).

A cluster analysis of the 33 differentially expressed proteins indicated that the classification accuracy was about 82% (Figure 2B). The volcano plot shows 11 proteins with differential expression based on the Bonferroni cutoff:  $p\text{-value} < 0.00018$ ,  $\log_2^{FC} > 1$  (Figure 2C). Compared to the control group, there were 4 upregulated and 7 downregulated proteins, including CEACAM-1 (carcinoembryonic antigen related cell adhesion molecule 1), GDF15 (growth differentiation factor 15), ACRP30 (Adiponectin), Nidogen-1, AXL (Tyrosine-protein kinase receptor UFO), FGF-9 (Fibroblast growth factor 9), Fas/TNFRSF6 (Apoptosis-mediating surface antigen FAS/Tumor necrosis factor receptor superfamily member 6), IGFBP-3 (Insulin-like growth factor-binding protein 3), Fc $\gamma$  RIIB/C (type IIB/C Fc receptor), MMP-9 (Matrix metalloproteinase 9), LYVE1 (lymphatic vessel endothelial hyaluronic acid receptor 1).

To elucidate the roles of 33 (intersection of three methods, Figure 2A) differentially expressed proteins in HCC patients, GO enrichment analysis was performed with a threshold of  $p < 0.05$ . The analysis was classified into three functional groups: cellular components (Figure 3A), molecular functions (Figure 3B), and biological processes (Figure 3C). In the cellular component group, aberrantly expressed proteins were mainly enriched in the extracellular region and plasma membrane. In the molecular function group, cytokine activity, growth factor activity, extracellular matrix binding, and chemokine activity were the main areas of enrichment. The results indicate that differential proteins are mainly enriched in biological processes of positive regulation of cell proliferation, cell death, cell motion and



**Figure 1.** Sample screening with semi-quantitative antibody array. 25 pairs of HCC and adjacent control tissue samples from HCC patients were screened using an antibody array (AAH-CYT-G4000). Changes in expression levels in 274 proteins, including growth factors, inflammation factors, angiogenesis, apoptosis factors, and adhesion molecules were examined.

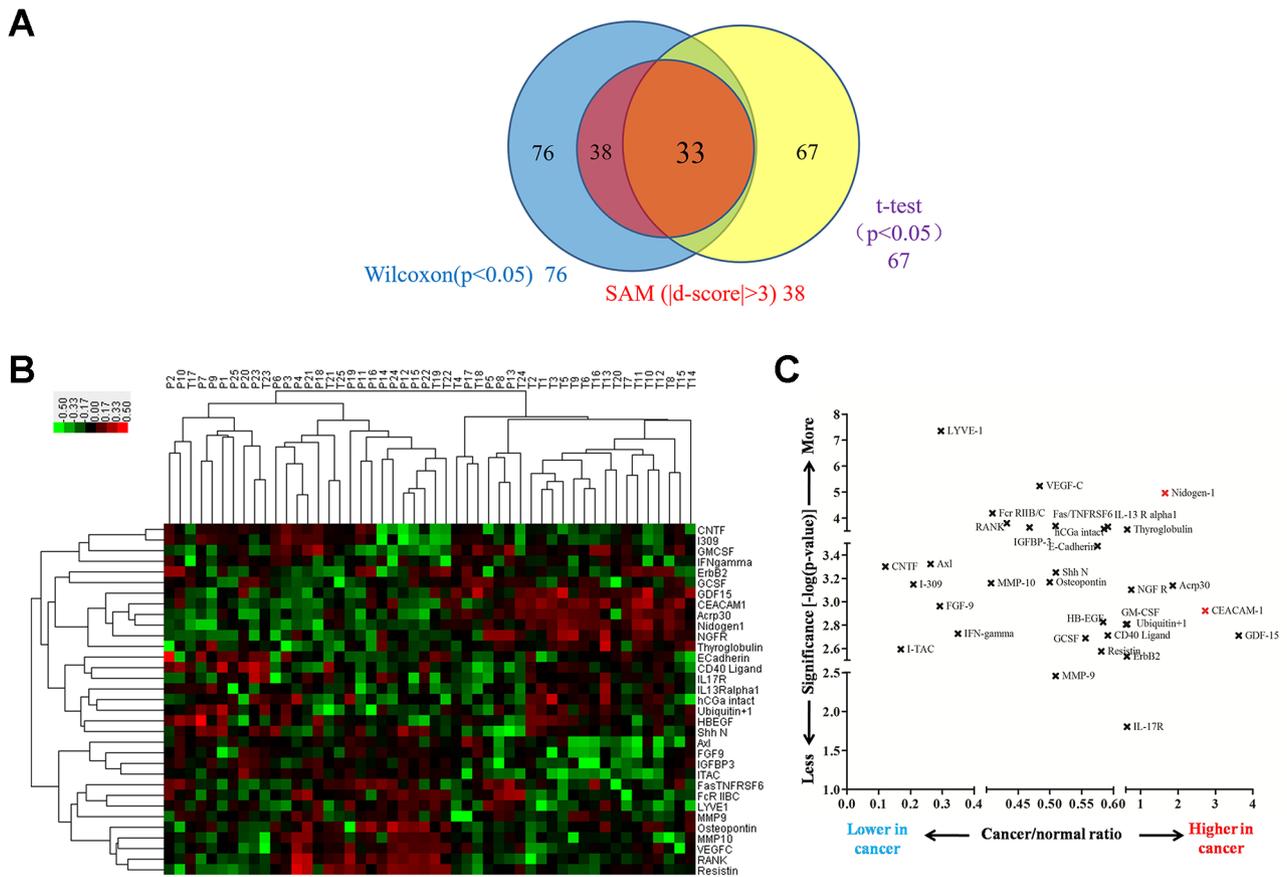


Figure 2. Protein expression analysis of 25 patients. A) Venn diagram for three different statistical analysis methods (t-test, Wilcoxon test, and SAM). B) Heatmaps for hierarchical clustering were obtained from differential proteins in HCC patient tissues and para-cancerous tissues. Highly expressed proteins in the tumor are shown in red, and those with low expression are shown in green. C) Volcano plot visualizing and identifying significantly differentially expressed proteins according to their log FC (x-axis: FC: fold change) and significance (y-axis:  $-\log_{10}$  adjusted p-value) in the HCC and control groups.

Table 1. Basic information of tissue samples from 25 HCC patients.

		HCC (n=25)
Gender	Male	21
	Female	4
Age	<55	15
	≥55	10
BCLC stage	A	18
	B	2
	C	5
	D	0

cytokine-mediated signaling pathway, and regulation of cell adhesion.

KEGG analysis further explored which signaling pathways are predominantly enriched in the differentially expressed proteins. According to KEGG enrichment analysis, the differentially expressed proteins were primarily enriched in cytokine-cytokine receptor interaction, autoimmune thyroid disease, JAK-STAT signaling pathway, allograft rejection,

bladder cancer, and pathways in cancer (Figure 3D). The results of GO and KEGG enrichment analyses suggest that the differentially expressed proteins are closely related to cell apoptosis, cell growth, and cell adhesion.

**Serum sample validation of 21 tumor markers and model construction.** According to the protein marker screening results of the HCC tissues and adjacent tissues, we identified 11 differentially expressed proteins. To further identify more effective biomarkers in the detection of HCC, we combined the 11 differentially expressed proteins identified in the tissue samples with 10 commonly used serum biomarkers previously reported in the literature to verify with a large number of serum samples from normal healthy controls and HCC cancer patients. A total of 21 biomarkers (Table 3) were measured in 76 HCC patients and 120 control samples (80 healthy, 10 esophageal cancer, 10 lung cancer, and 20 gastric cancer) using a custom antibody array (Raybiotech, Inc). 18 biomarkers showed significant differential expression between the two groups (FDR < 0.05, Table 3). The Yeo-Johnson transformation is an

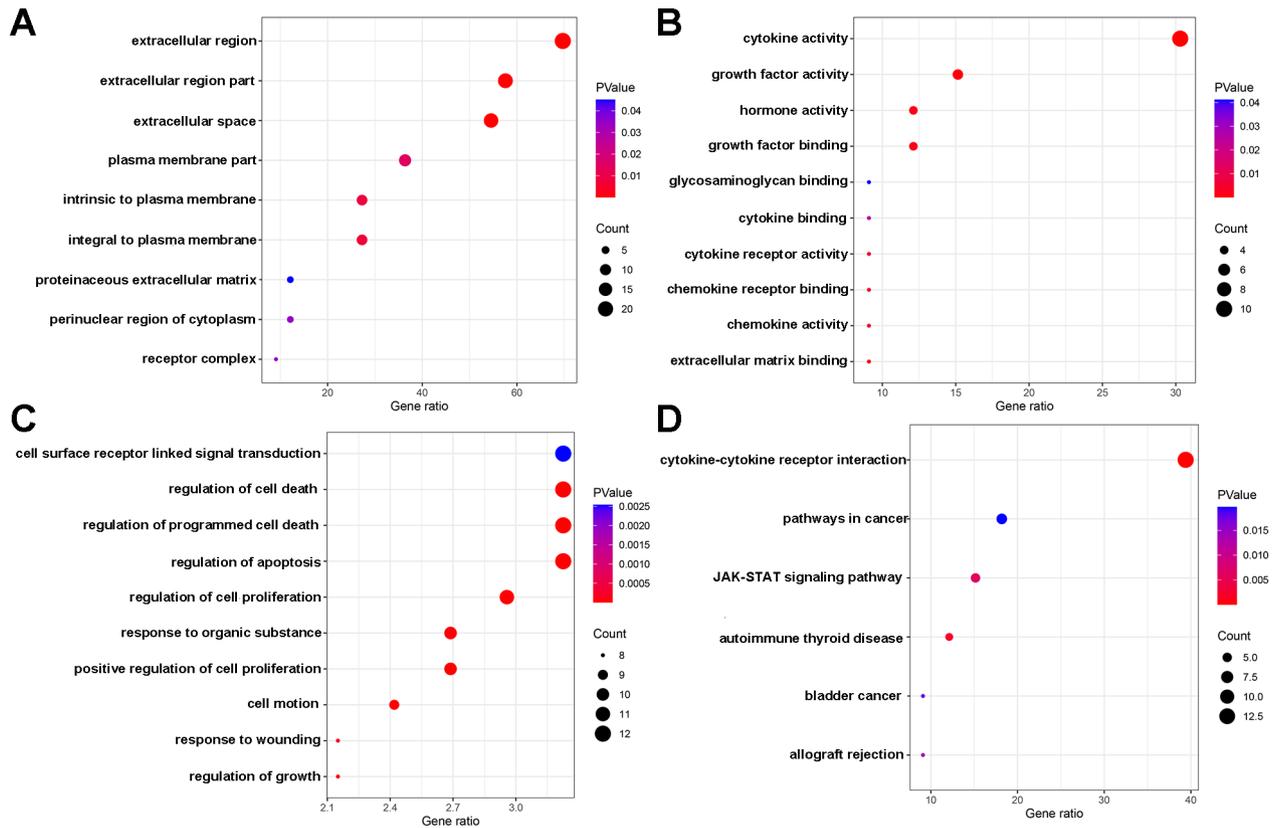


Figure 3. Gene ontology and KEGG analyses in the differentially expressed proteins.  $p < 0.05$  was considered statistically significant. Gene ontology enrichment analysis includes three subtypes: cellular component (A), molecular function (B), and biological process (C). D) KEGG pathway enrichments in the differentially expressed proteins.

Table 2. Basic information of serum samples from 76 HCC patients and 120 controls.

		HCC	Control (n=120)			
		(n=76)	Lung cancer (n=10)	Esophageal cancer (n=10)	Gastric cancer (n=20)	Normal (n=80)
Gender	Male	69	9	6	16	53
	Female	7	1	4	4	27
Age	<55	39	2	0	5	53
	$\geq 55$	37	8	10	15	27

extension of the Box-Cox transformation, which allows the transformation of data containing zero or negative values. All the following analyses, including model construction and collinearity analysis, were implemented with the transformed 21 biomarker values. Herein, we excluded 3 biomarkers: Fcy RIIB/C, VEGF, and SCCA1, that were not sufficiently different ( $FDR \geq 0.05$ ) and 1 biomarker that was highly collinear ( $TGF-\beta 1$ , average  $R > 0.7$ ) (Figure 4) from the 21 biomarkers to facilitate subsequent model building with 17 differential proteins.

The 17 biomarkers: AFP, GDF-15, CEACAM-1, MMP-9, GP73, B2M, IGFBP-3, ACRP30, Ferritin, Axl, LYVE-1, Fas, DKK-1, HGF, IL-8, FGF-9, and Nidogen-1 were used

to construct 4 different models, logistic regression (LR), linear discriminant analysis (LDA), random forest (RF), and support vector machine (SVM). The models trained with the 17 biomarkers demonstrated good performance (Figures 5A, 5B). Models were trained with a different number of biomarkers from 16 to 2. We found that the model performance with 4 biomarkers was as good as that of 17 biomarkers (Figures 5C, 5D). The four models constructed with different numbers of protein biomarkers may have their own advantages in the prediction accuracy of the test set (Figure 5E). Therefore, multiple models were constructed and compared, rather than a single model, and the optimal model RF was selected.

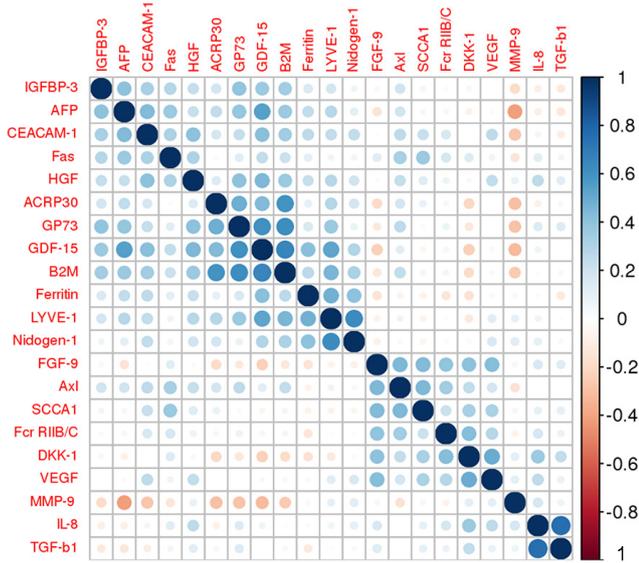


Figure 4. Collinearity between the 21 biomarkers after Yeo-Johnson transformation.

Figure 5A shows the ROC curves of the four 17-biomarker models using the training dataset. The accuracy and kappa values, which are used to evaluate the model built on the training dataset with the test dataset are shown in Figure 5B. The method of 17 protein biomarkers to assess HCC is still complex in practice, 17 is still a large number. Therefore, it is necessary to find protein biomarkers with a number smaller than 17 in the process of dimensionality reduction, while maintaining the accuracy of the classification and prediction of the model.

As the dimensionality reduction (reducing the biomarker number from 16 to 2) model construction proceeded, it was found that four protein markers: AFP, GDF-15, CEACAM-1, and MMP-9 could be well constructed for model building and still maintain high prediction accuracy (Figure 5E).

A comparison of the accuracy of the four models demonstrated that RF is a method that can ensure high classification accuracy of both the training dataset and the test dataset (Figures 5F, 5G). When 17 protein markers are modeled by RF, the accuracy of the training dataset and test dataset are 1 and 0.949 (Table 4), respectively. The accuracy of the training

Table 3. Comparison of expression levels of biomarkers between the control and HCC groups.

Variable	Control	HCC	Statistic	p-value	FDR
AFP	1033.32 (78.65; 6018.6)	4457.6 (207.75; 11439.2)	W=950	0	0
GDF-15	575.87 (164.96; 1980.85)	1479.86 (296.2; 2688.43)	W=1255	0	0
CEACAM-1	1185.46 (341.5; 23978.71)	2271.04 (287.02; 10145.78)	W=1650	0	0
MMP-9	4579.14 (0; 12715.81)	2002.07 (344.02; 7820.54)	W=7503	0	0
GP73	6075.82 (530.93; 15963.3)	10287.03 (2651.19; 21120.25)	W=1765	0	0
B2M	644.31 (291.56; 2377.27)	1152.6 (374.78; 2501)	W=1818	0	0
IGFBP-3	51627.25 (6108.52; 97901.45)	65636.13 (20925.56; 174243.1)	W=2381	0	0
ACRP30	3322.69 (159.75; 47097.56)	10871.04 (273.28; 189290.9)	W=2736	0.0000008	0.0000021
Ferritin	3556.1 (4.25; 16876.18)	8175.08 (114.5; 25044.08)	W=2713	0.0000018	0.0000043
Axl	5.78 (0; 3536.73)	10.72 (0.38; 72.54)	W=2916	0.0000076	0.000016
LYVE-1	1540.6 (0; 3237.66)	2089.93 (887.84; 5868.59)	W=2846	0.0000095	0.0000181
TGF-β1	1497.25 (0; 7339.73)	826.99 (0; 3478.18)	W=6238	0.000077	0.0001348
Fas	10.86 (0.23; 1961.94)	16.2 (2.71; 1890.45)	W=3210	0.0001916	0.0003095
DKK-1	53.92 (0; 3655.42)	26.03 (0; 639.7)	W=6091.5	0.0003405	0.0005108
HGF	245.02 (77.68; 2083.29)	337.27 (64.13; 1888.87)	W=3400	0.0011637	0.0016292
IL-8	67.89 (0; 341.43)	46.45 (0; 166.04)	W=5874	0.0024508	0.0032167
FGF-9	80.69 (0; 1178.2)	65.83 (0; 441.59)	W=5741	0.0070823	0.0087487
Nidogen-1	6018.17 (0.26; 14229.39)	6797.5 (1202.72; 15365.03)	W=3713	0.0286917	0.0334737
SCCA1	346.79 (0; 18681.44)	306.56 (0; 1833.73)	W=4982.5	0.4433278	0.4726889
VEGF	93.61 (10.25; 3746.45)	89.1 (2.74; 738.89)	W=4978	0.4501799	0.4726889
Fcγ. RIIB/C	234.1 (0; 3535.17)	214.32 (0; 2066.69)	W=4906.5	0.5662143	0.5662143

Table 4. Performance of 17-biomarker models using the training dataset.

	17-biomarker models in training set				17-biomarker models in test set				
	threshold	sensitivity	specificity	accuracy	threshold	sensitivity	specificity	accuracy	
LR	0.375	0.967	0.948	0.955	LR	0.478	0.933	0.875	0.897
LDA	0.294	0.918	0.958	0.943	LDA	0.568	0.933	0.917	0.923
RF	0.483	1.000	1.000	1.000	RF	0.617	0.933	0.958	0.949
SVM	0.279	0.967	0.969	0.968	SVM	0.531	0.933	0.917	0.923

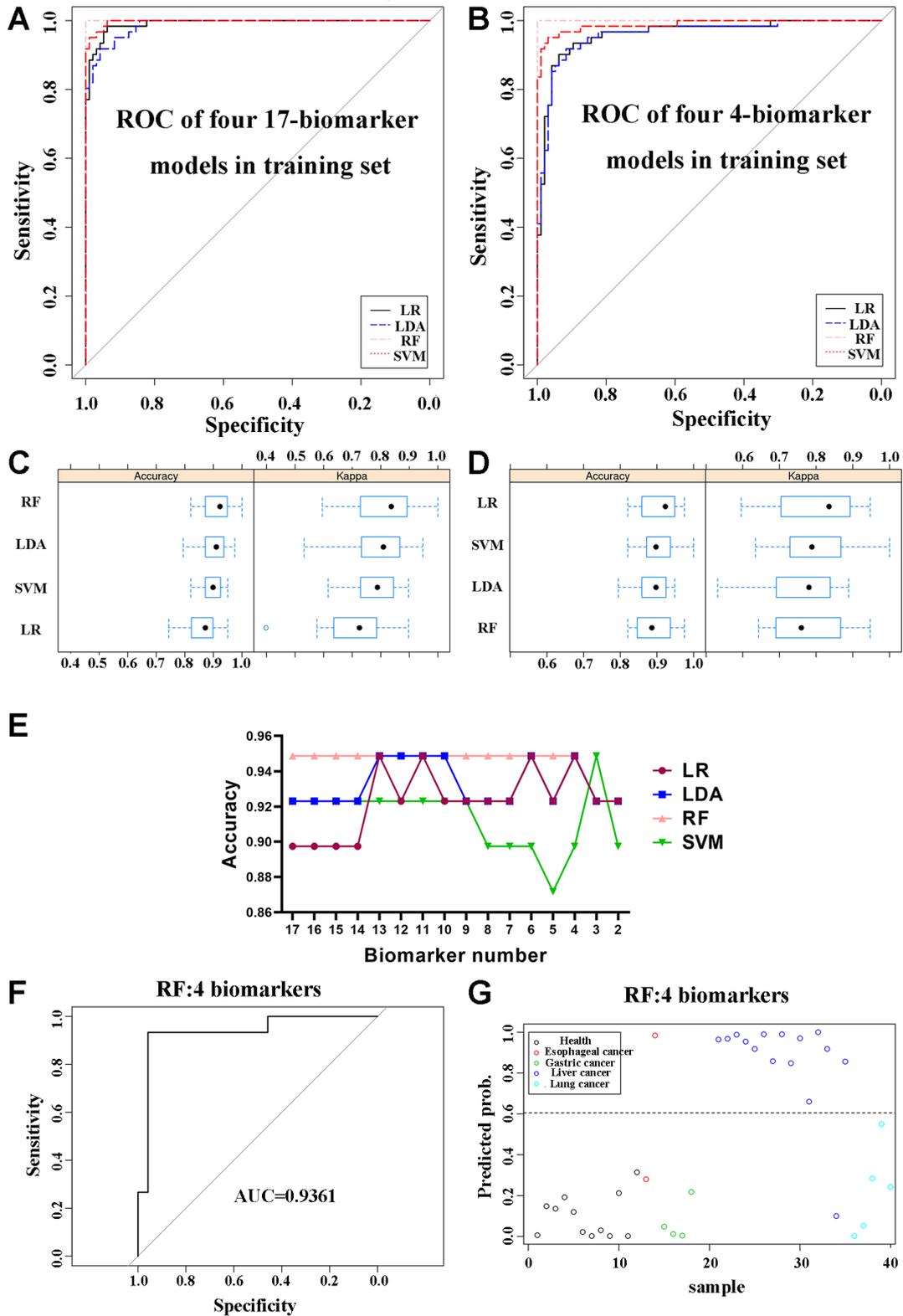


Figure 5. Model building and biomarker dimensionality reduction. A) ROC curves of four 17-biomarker models using the training dataset. B) Cross-validation accuracy and kappa values of the four 17-biomarker models. C) ROC curves of four 4-biomarker models using the training dataset. D) Cross-validation accuracy and kappa values of the four 4-biomarker models. E) Accuracy of the four models using different numbers of biomarkers. F) The 4-biomarker RF model prediction using the training dataset. G) The 4-biomarker RF model prediction using the test dataset.

**Table 5. Performance of 4-biomarker models using the training dataset.**

	4-biomarker models in training set				4-biomarker models in test set				
	threshold	sensitivity	specificity	accuracy	threshold	sensitivity	specificity	accuracy	
LR	0.382	0.902	0.938	0.924	LR	0.845	0.933	0.958	0.949
LDA	0.220	0.918	0.917	0.917	LDA	0.779	0.933	0.958	0.949
RF	0.502	1.000	1.000	1.000	RF	0.605	0.933	0.958	0.949
SVM	0.592	0.951	0.969	0.962	SVM	0.606	0.933	0.875	0.897

and test datasets is the same as the model for the four protein markers (Table 5). Therefore, we identified three novel protein biomarkers (GDF-15, CEACAM-1, and MMP-9) from HCC tissue that can be utilized non-invasively in serum in addition to the well-characterized AFP with high accuracy to differentiate HCC patients from controls.

## Discussion

To identify effective biomarkers for HCC, 11 differentially expressed proteins were discovered in HCC tissues and paracancerous tissues of 25 HCC patients. Since the differentially expressed proteins found at the tissue level may not necessarily be different in serum, verification of a large number of serum samples is required. Therefore, 11 differential proteins found in tissues were combined with 10 protein markers commonly used in HCC serum detection reported in the literature and validated with a large number of serum samples.

Through the verification of serum samples, 4 protein biomarkers (AFP, GDF-15, CEACAM-1, and MMP-9) were finally selected, and HCC was well predicted in the test dataset by random forest. Among them, AFP is a commonly used clinical diagnostic biomarker for HCC. GDF-15 and CEACAM-1 were highly expressed and MMP-9 was weakly expressed in HCC patients compared to control patients. Wang et al. found that GDF15 is positively associated with the elevation of Treg cell frequencies in patients with HCC. The study also noted gene ablation of GDF15 in HCC can convert an immunosuppressive tumor microenvironment to an inflammatory state. Generation and function enhancement of Treg cells induced by GDF15 is a new mechanism for HCC-related immunosuppression [17]. The expression of GDF15 was significantly upregulated in HCC cells exposed to chemotherapeutic agents. GDF15 from chemotherapy-damaged HCC cells promoted the *in vitro* proliferation, migration, and tube formation of endothelial cells. The pro-angiogenic effect of GDF15 was through the activation of Src and its downstream AKT, MAPK, and NF- $\kappa$ B signaling [18]. Serum GDF15 is positively related to the levels of PIVKA-II and AFP in patients with HCC and GDF15 is a potent serum marker for the detection of HBV-associated HCC. PIVKA-II combined with GDF15 was found to improve diagnostic accuracy for HBV-associated HCC [19]. GDF15 is a protein closely related to HCC and can be used in combination with other markers to improve diagnostic accuracy. CEACAM-1

is a transmembrane glycoprotein and a member of the carcinoembryonic antigen family. CEACAM-1 expression was positively correlated with the expression of EMT-related factors and microvessel density of tumor tissues in HCC [20]. The mechanism of action of CEACAM-1 in tumors is not very clear, and it was thought to inhibit tumorigenesis in the early stage. However, some studies [20, 21] have shown that it promotes angiogenesis and lymphangiogenesis, thereby promoting tumor growth. Herein, the detected CEACAM-1 level in the HCC tissue and serum was higher than that of the control group. MMP-9 is considered to be a matrix metalloproteinase closely related to tumor invasion and metastasis. A recent study found that the expression of MMP-9 in HCC has a certain relationship with tumor invasion and metastasis [22–24]. However, to date, the research on MMP-9 in HCC is relatively limited, and the mechanism is not very clear. It is possible that the interaction of MMP-9 with CD4<sup>+</sup>, CD8<sup>+</sup>, TNF- $\alpha$ , etc. may be one of the mechanisms affecting the invasion and metastasis of HCC. Most studies have shown that MMP-9 is overexpressed in HCC [25–29], but controversially, our experimental results show that MMP-9 is decreased in both serum and tissue compared to healthy controls. Most of the selected samples may be in the A and B stages of BCLC staging and have not yet experienced metastasis and invasion, which may explain the downregulation of MMP-9 in the HCC group. Further studies with additional samples at different stages would be required to determine the role of MMP-9 in HCC more clearly.

In conclusion, we found four protein biomarkers that can monitor HCC in serum. The combination of the four proteins can well avoid the problems that may arise when AFP is used alone as a diagnostic basis, and greatly improves the sensitivity, specificity, and accuracy of HCC diagnosis, thus it is probably to be a powerful tool for the clinical diagnosis of HCC.

**Supplementary information** is available in the online version of the paper.

**Acknowledgments:** This research was funded by the following grants: RayBiotech innovative research fund, Guangzhou Innovation Leadership Team (CXLJTD-201602), Pearl River S&T Nova Program of Guangzhou (201806010035, 201806010032), Guangdong Province Key Technologies R&D Program for “Precision Medicine and Stem Cells” (2019B020231002,

2019B020227004), National Science and Technology Major Project of China (No.2018ZX10732-401-003-012), the Industry-University-Research Collaborative Innovation Special Project of Guangzhou (201802030001). Guangzhou Basic and Applied Basic Research Project (202002030242), Basic Research Plan of Guangzhou People's Livelihood Science and Technology Project (202002020021), and Guangzhou key medical discipline construction project fund, the National Key R&D Program of China (2018YFC1106300, 2017YFC1105000), the Joint Funds of the National Natural Science Foundation of China (No. U1501245), the National Natural Science Foundation of China (No.51672088). National Natural Science Foundation of China (82173149), Basic and applied basic research funding Guangdong Province (2021A1515012441), The Science and Technology Program of Huizhou Daya Bay (2021002). Shuangzhe Zhang, Gordon F. Huang, Hao Tang, Zhuo Zhang, Xuedong Song, Kelly Whittaker, and Ruo-Pan Huang are employees of RayBiotech and have a financial stake in RayBiotech.

## References

- [1] XU J. Trends in Liver Cancer Mortality Among Adults Aged 25 and Over in the United States, 2000–2016. *NCHS Data Brief* 2018; 314: 1–8.
- [2] JACOB JA. Liver Cancer Death Rate Increases as Overall Cancer Deaths Decline. *JAMA* 2016; 315: 1692. doi:10.1001/jama.2016.3599
- [3] WANG JJ, WANG LH, ZHU YE, ZHANG JY, LIAO J et al. A high accuracy cantilever array sensor for early liver cancer diagnosis. *Biomed Microdevices* 2016; 18: 110. <https://doi.org/10.1007/s10544-016-0132-5>
- [4] JOO I, CHOI BI. New Paradigm for Management of Hepatocellular Carcinoma by Imaging. *Liver Cancer* 2012; 1: 94–109. <https://doi.org/10.1159/000342404>
- [5] PARK J, LEE JM, KIM TH, YOON JH. Imaging Diagnosis of HCC: Future directions with special emphasis on hepatobiliary MRI and contrast-enhanced ultrasound. *Clin Mol Hepatol* 2022; 28: 362–379. <https://doi.org/10.3350/cmh.2021.0361>
- [6] WINKENS T, RUDAKOFF W, RAUCHFUSS F, MALESSA C, SETTMACHER U et al. FDG PET/CT to Detect Incidental Findings in Patients With Hepatocellular Carcinoma-Additional Benefit for Patients Considered for Liver Transplantation? *Clin Nucl Med* 2021; 46: 532–539. <https://doi.org/10.1097/RLU.0000000000003576>
- [7] BARGELLINI I, BATTAGLIA V, BOZZI E, LAURETTI DL, LORENZONI G et al. Radiological diagnosis of hepatocellular carcinoma. *J Hepatocell Carcinoma* 2014; 1: 137–148. <https://doi.org/10.2147/JHC.S44379>
- [8] ZHANG ZG, ZHANG YY, WANG YY, XU LL, XU WJ. Alpha-fetoprotein-L3 and Golgi protein 73 may serve as candidate biomarkers for diagnosing alpha-fetoprotein-negative hepatocellular carcinoma. *Onco Targets Ther* 2015; 9: 123–129. <https://doi.org/10.2147/OTT.S90732>
- [9] ZHAO YJ, JU Q, LI GC. Tumor markers for hepatocellular carcinoma. *Mol Clin Oncol* 2013; 1: 593–598. <https://doi.org/10.3892/mco.2013.119>
- [10] AN J, KIM HI, CHANG S, SHIM JH. Continued value of the serum alpha-fetoprotein test in surveilling at-risk populations for hepatocellular carcinoma. *PLoS One* 2020; 15: e0238078. <https://doi.org/10.1371/journal.pone.0238078>
- [11] TANAKA T, TANIGUCHI T, SANNOMIYA K, TAKENAKA H, TOMONARI T et al. Novel des--carboxy prothrombin in serum for the diagnosis of hepatocellular carcinoma. *J Gastroenterol Hepatol* 2013; 28: 1348–1355. <https://doi.org/10.1111/jgh.12166>
- [12] ATTALLAH AM, EL-FAR M, MALAK CAA, OMRAN MM, FARID K et al. A simple diagnostic index comprising epithelial membrane antigen and fibronectin for hepatocellular carcinoma. *Ann Hepatol* 2015; 14: 869–880. <https://doi.org/10.5604/16652681.1171774>
- [13] ZHANG P, XIAO J, RUAN Y, ZHANG Z, ZHANG X. Monitoring Value of Serum HER2 as a Predictive Biomarker in Patients with Metastatic Breast Cancer. *Cancer Manag Res* 2020; 12: 4667–4675. <https://doi.org/10.2147/CMAR.S254897>
- [14] CHAUHAN R, LAHIRI N. Tissue- and Serum-Associated Biomarkers of Hepatocellular Carcinoma. *Biomark Cancer* 2016; 8: 37–55. <https://doi.org/10.4137/BIC.S34413>
- [15] LOZADA ME, CHAITEERAKIJ R, ROBERTS LR. Screening for hepatocellular carcinoma and cholangiocarcinoma: Can biomarkers replace imaging? *Curr Hepatol Rep* 2015; 14: 128–138. <https://doi.org/10.1007/s11901-015-0261-y>
- [16] ZHOU Q, MAO YQ, JIANG WD, CHEN YR, HUANG RY et al. Development of IGF Signaling Antibody Arrays for the Identification of Hepatocellular Carcinoma Biomarkers. *PLoS One* 2012; 7: e46851. <https://doi.org/10.1371/journal.pone.0046851>
- [17] WANG Z, HE L, LI W, XU C, ZHANG J et al. GDF15 induces immunosuppression via CD48 on regulatory T cells in hepatocellular carcinoma. *J Immunother Cancer* 2021; 9: e002787. <https://doi.org/10.1136/jitc-2021-002787>
- [18] DONG G, ZHENG QD, MA M, WU SF, ZHANG R et al. Angiogenesis enhanced by treatment damage to hepatocellular carcinoma through the release of GDF15. *Cancer Med* 2018; 7: 820–830. <https://doi.org/10.1002/cam4.1330>
- [19] CHEN J, TANG D, XU C, NIU Z, LI H et al. Evaluation of Serum GDF15, AFP, and PIVKA-II as Diagnostic Markers for HBV-Associated Hepatocellular Carcinoma. *Lab Med* 2021; 52: 381–389. <https://doi.org/10.1093/labmed/lmaa089>
- [20] YOSHIKAWA M, MORINE Y, IKEMOTO T, IMURA S, HIGASHIJIMA JUN et al. Elevated Preoperative Serum CEA Level Is Associated with Poor Prognosis in Patients with Hepatocellular Carcinoma Through the Epithelial–Mesenchymal Transition. *Anticancer Res* 2017; 37: 1169–1175. <https://doi.org/10.21873/anticancer.11430>
- [21] SERRA S, ASA SL, BAMBERGER AM, WAGENER C, CHETTY R. CEACAM1 Expression in Pancreatic Endocrine Tumors. *Appl Immunohistochem Mol Morphol* 2009; 17: 286–293. <https://doi.org/10.1097/PAI.0b013e318196e13c>
- [22] ILLEMANN M, BIRD N, MAJEED A, SEHESTED M, LAERUM OD et al. MMP-9 Is Differentially Expressed in Primary Human Colorectal Adenocarcinomas and Their Metastases. *Mol Cancer Res* 2006; 4: 293–302. <https://doi.org/10.1158/1541-7786.MCR-06-0003>

- [23] SAKAMOTO Y, MAFUNE K, MORI M, SHIRAISHI T, IMAMURA H et al. Overexpression of MMP-9 correlates with growth of small hepatocellular carcinoma. *Int J Oncol* 2000; 17: 237–243. <https://doi.org/10.3892/ijo.17.2.237>
- [24] ZENG ZS, GUILLEM JG. Unique activation of matrix metalloproteinase-9 within human liver metastasis from colorectal cancer. *Br J Cancer* 1998; 78: 349–353. <https://doi.org/10.1038/bjc.1998.497>
- [25] EISA NH, EBRAHIM MA, RAGAB M, EISSA LA, EL-GAYAR AM. Galectin-3 and matrix metalloproteinase-9: perspective in management of hepatocellular carcinoma. *J Oncol Pharm Pract* 2015; 21: 323–330. <https://doi.org/10.1177/1078155214532698>
- [26] CHEN JS, HUANG XH, WANG Q, HUANG JQ, ZHANG LJ et al. Sonic hedgehog signaling pathway induces cell migration and invasion through focal adhesion kinase/AKT signaling-mediated activation of matrix metalloproteinase (MMP)-2 and MMP-9 in liver cancer. *Carcinogenesis* 2012; 34: 10–19. <https://doi.org/10.1093/carcin/bgs274>
- [27] ROOMI MW, MONTERREY JC, KALINOVSKY T, RATH M, NIEDZWIECKI A. Patterns of MMP-2 and MMP-9 expression in human cancer cell lines. *Oncol Rep* 2009; 21: 1323–1333. [https://doi.org/10.3892/or\\_00000358](https://doi.org/10.3892/or_00000358)
- [28] Ordonez R, Carbajo-Pescador S, Prieto-Dominguez N, Garcia-Palomo A, Gonzalez-Gallego J et al. Inhibition of matrix metalloproteinase-9 and nuclear factor kappa B contribute to melatonin prevention of motility and invasiveness in HepG2 liver cancer cells. *J Pineal Res* 2014; 56: 20–30. <https://doi.org/10.1111/jpi.12092>
- [29] WANG J, LIU Q, XIAO H, LUO X, LIU X. Suppressive effects of Momordin Ic on HepG2 cell migration and invasion by regulating MMP-9 and adhesion molecules: Involvement of p38 and JNK pathways. *Toxicol In Vitro* 2019; 56: 75–83. <https://doi.org/10.1016/j.tiv.2019.01.007>

[https://doi.org/10.4149/neo\\_2022\\_220606N600](https://doi.org/10.4149/neo_2022_220606N600)

## Identification and validation of circulating biomarkers for detection of liver cancer with antibody array

Shuangzhe ZHANG<sup>1,2,3,\*</sup>, Chunhui GAO<sup>4,\*</sup>, Qi ZHOU<sup>5,6,\*</sup>, Lipai CHEN<sup>4,\*</sup>, Wei HUANG<sup>1</sup>, Gordon Fan HUANG<sup>7</sup>, Hao TANG<sup>1,7</sup>, Xuedong SONG<sup>1</sup>, Zhuo ZHANG<sup>1</sup>, Kelly WHITTAKER<sup>7</sup>, Xiaofeng CHEN<sup>3,8,9</sup>, Ruo-Pan HUANG<sup>1,2,7,\*</sup>

### Supplementary Information

#### Supplementary Table S1. 274 proteins of AAH-CYT-G4000.

4-1BB (TNFRSF9/CD137)	ACE-2	Adiponectin (ACRP30)	Activin A	Adipsin (Complement Factor D)
AgRP	ALCAM (CD166)	Alpha-fetoprotein	Amphiregulin	Angiogenin
Angiopoietin-1	Angiopoietin-2	ANGPTL4	Axl	
CD80 (B7-1)	Beta-2 Microglobulin	BCAM	BCMA (TNFRSF17)	BDNF
beta IG-H3	bFGF	BLC (CXCL13)	BMP-4	BMP-5
BMP-6	BMP-7	beta-NGF	Betacellulin (BTC)	CA125
CA15-3	CA19-9	CA9	Cardiotrophin-1 (CT-1)	Cathepsin S
HCC-1 (CCL14)	6Ckine (CCL21)	CCL28 (MEC)	CD14	CD23
CD30 (TNFRSF8)	CD40 (TNFRSF5)	CD40 Ligand (TNFSF5)	CEA	CEACAM-1
CK beta 8-1 (CCL23)	CNTF	Cripto-1	CRP (C-Reactive Protein)	CTACK (CCL27)
CXCL16	DAN	Decorin	DKK-1	Dkk-3
Dkk-4	CD26 (DPPIV)	DR6 (TNFRSF21)	Dtk	E-Cadherin
EDA-A2	EGF	EGFR	EG-VEGF (PK1)	ENA-78 (CXCL5)
Endoglin (CD105)	Eotaxin-1 (CCL11)	Eotaxin-2 (MPIF-2/CCL24)	Eotaxin-3 (CCL26)	TROP1 (EpCAM)
ErbB2	ErbB3	Erythropoietin R	E-Selectin	Fas (TNFRSF6/Apo-1)
Fas Ligand (TNFSF6)	Fc gamma RIIB/C (CD32B/C)	Ferritin	FGF-4	FGF-6
FGF-7 (KGF)	FGF-9	Flt-3 Ligand	FLRG	Follistatin
Fractalkine (CX3CL1)	FSH	Furin	Galectin-7	GCP-2 (CXCL6)
GCSF	GDF-15	GDNF	GITR (TNFRSF18)	GITR Ligand (TNFSF18)
GM-CSF	GRO alpha/beta/gamma	GRO alpha (CXCL1)	Growth Hormone	HB-EGF
HCC-4 (CCL16)	hCG intact	HGF	HVEM (TNFRSF14)	I-309 (TCA-3/CCL1)
ICAM-1 (CD54)	ICAM-2 (CD102)	ICAM-3 (CD50)	IFN-gamma	IGFBP-1
IGFBP-2	IGFBP-3	IGFBP-4	IGFBP-6	IGF-1
IGF-1 R	IGF-2	IL-1 R2	IL-1 R4 (ST2)	IL-1 R1

Supplementary Table S1. *Continued . . .*

IL-10	IL-10 R alpha	IL-10 R beta	IL-11	IL-12 p40
IL-12 p70	IL-13	IL-13 R alpha 2	IL-13 R1	IL-15
IL-16	IL-17A	IL-17B	IL-17C	IL-17F
IL-17 RA	IL-18 BP alpha	IL-18 R beta (AcPL)	IL-1 alpha (IL-1 F1)	IL-1 beta (IL-1 F2)
IL-1 ra (IL-1 F3)	IL-2	IL-2 R beta (CD122)	IL-2 R gamma (Common gamma Chain)	IL-2 R alpha
IL-21 R	IL-22	IL-28A (IFN-lambda 2)	IL-29 (IFN-lambda 1)	IL-3
IL-31	IL-4	IL-5	IL-5 R alpha	IL-6
IL-6 R	IL-7	IL-8 (CXCL8)	IL-9	Insulin
IP-10 (CXCL10)	I-TAC (CXCL11)	LAP/TGF beta 1	Leptin	Leptin R
LIF	Light (TNFSF14)	LIMPII	L-Selectin (CD62L)	Luteinizing hormone
Lymphotoctin (XCL1)	LYVE-1	Marapsin	MCP-1 (CCL2)	MCP-2 (CCL8)
MCP-3 (MARC/CCL7)	MCP-4 (CCL13)	M-CSF	M-CSF R	MDC (CCL22)
MICA	MICB	MIF	MIG (CXCL9)	MIP-1 alpha (CCL3)
MIP-1 beta (CCL4)	MIP-1 delta (CCL15)	MIP-3 alpha (CCL20)	MIP-3 beta (CCL19)	MMP-1
MMP-10	MMP-13	MMP-2	MMP-3	MMP-7
MMP-8	MMP-9	MPIF-1 (CCL23)	MSP alpha/beta	NAP-2 (PPBP/CXCL7)
NCAM-1 (CD56)	NGFR (TNFRSF16)	Nidogen-1	NrCAM	NRG1-beta 1 (HRG1-beta 1)
NT-3	NT-4	Oncostatin M	Osteopontin (SPP1)	Osteoprotegerin (TNFRSF11B)
PAI-1	PARC (CCL18)	PDGF-AA	PDGF R alpha	PDGF R beta
PDGF-AB	PDGF-BB	PECAM-1 (CD31)	PLGF	Platelet factor 4 (CXCL4)
Procalcitonin	Prolactin	PSA-free	PSA-total	RAGE
RANK (TNFRSF11A)	RANTES (CCL5)	Resistin	S100 B	SAA (Serum Amyloid A)
SCF	SCF R (CD117/c-kit)	SDF-1 alpha (CXCL12 alpha)	SDF-1 beta (CXCL12 beta)	gp130
Sonic Hedgehog N- Terminal (Shh-N)	Siglec-5 (CD170)	Siglec-9	TNF RII (TNFRSF1B)	TNF RI (TNFRSF1A)
TACE	TARC (CCL17)	TECK (CCL25)	TGF beta 2	TGF alpha
TGF beta 3	TGF beta 1	Thrombopoietin (TPO)	Thyroglobulin	Tie-1
Tie-2	TIM-1 (KIM-1)	TIMP-1	TIMP-2	TIMP-4
TNF alpha	TNF beta (TNFSF1B)	TRAIL R2 (TNFRSF10B/DR5)	TRAIL R3 (TNFRSF10C)	TRAIL R4 (TNFRSF10D)
Trappin-2	TREM-1	TSH	TSLP	Ubiquitin+1
uPAR	VCAM-1 (CD106)	VE-Cadherin (CDH5)	VEGF-A	VEGFR2
VEGFR3	VEGF-C	VEGF-D	XEDAR	

**Supplementary Table S2. SAM (Significant Analysis of Microarray, FDR < 5%) screening of 69 differentially expressed proteins between HCC tissues and para-cancer tissues.**

Protein	d-score	Protein	d-score	Protein	d-score
Acrp30	4.040771312	Follistatin	2.017659974	I-TAC	-2.642609461
Angiogenin	2.119673499	GCSF	-3.152900461	LAP	1.425004374
Axl	-2.25864129	GDF-15	3.360832432	LYVE-1	-7.524597203
BCAM	2.517766815	GDNF	-1.769272479	MCP-4	-1.588191178
BDNF	-2.258750686	GM-CSF	-2.694724607	M-CSF	-1.829620965
BMP-6	-1.824564228	HB-EGF	-1.986716928	MIP-3-alpha	-1.179841296
CD23	-1.418256489	hCGa, intact	-3.107843916	MMP-10	-1.911451919
CD40 Ligand	-2.97722624	HGF	-2.034775196	MMP-7	-1.682205813
CEACAM-1	3.456695869	I-309	-2.53637464	MMP-9	-3.096942344
CK beta 8-1	-1.575450545	IFN-gamma	-2.331008596	NGF R	-1.962834504
CNTF	-2.85163759	IGFBP-3	-2.602695507	Nidogen-1	5.260659198
CRP	1.842492184	IL-1 R4/ST2	-2.268898856	Osteopontin	-1.857941272
DKK-3	-1.474679665	IL-10	-2.105873245	PAI-I	2.348633003
DPPIV	1.833307786	IL-13	-1.87026642	RANK	-2.017881501
E-Cadherin	-3.441371544	IL-13 R alpha1	-2.326993768	Resistin	-2.091681873
EpCAM	2.188631011	IL-15	-2.2756541	Shh N	-2.636357956
ErbB2	-2.177143026	IL-17B	-1.452857553	Siglec-9	-1.39366937
ErbB3	2.071093038	IL-17R	-2.816358916	TGF-beta 1	-2.151464751
Fas/TNFRSF6	-4.159939102	IL-2	-2.404221092	Thrombopoietin	-2.089462051
Fcr RIIB/C	-3.361004704	IL-2 Rapha	-2.025754721	Thyroglobulin	-2.004136217
Ferritin	1.765537687	IL-3	-2.762248395	TIMP-1	2.093905379
FGF-4	-1.732302617	IL-5	-1.087064232	Ubiquitin+1	-1.765059908
FGF-9	-2.391656413	IL-7	-1.904175188	VEGF-C	-2.377831176

**Supplementary Table S3. A Wilcoxon test (p < 0.05) screening of 76 differentially expressed proteins between HCC tissues and para-cancer tissues.**

Protein	Wilcoxon test	Protein	Wilcoxon test	Protein	Wilcoxon test	Protein	Wilcoxon test
Acrp30	0.001	FGF-4	0.008	IL-17B	0.011	Osteopontin	0.001
Axl	0.003	FGF-9	0.002	IL-17R	0	PAI-I	0.002
BCAM	0.008	Follistatin	0.002	IL-1alpha	0	PDGF-BB	0.031
BDNF	0.002	GCSF	0.002	IL-2	0.002	RANK	0.001
BMP-6	0.007	GDF-15	0.001	IL-2 Rapha	0.005	Resistin	0.005
CA19-9	0.026	GDNF	0.003	IL-3	0.001	Shh N	0.001
CD14	0.011	GM-CSF	0	IL-5	0.02	Siglec-9	0.019
CD23	0.006	GRO-alpha	0.004	IL-7	0.012	TGF-beta 1	0.004
CD40 Ligand	0.001	HB-EGF	0.007	I-TAC	0	Thrombopoietin	0.005
CEA	0.045	hCGa, intact	0	LYVE-1	0	Thyrobulin	0
CEACAM-1	0.001	HGF	0.04	MCP-4	0.048	TIM-1	0.004
CK beta 8-1	0.009	HVEM	0.037	M-CSF	0.015	TIMP-1	0.042
CNTF	0.001	I-309	0.001	MIG	0.026	TNF-alpha	0.012
DKK-3	0.009	IFN-gamma	0.005	MIP-3-alpha	0.034	Ubiquitin+1	0.001
DPPIV	0.04	IGFBP-3	0	MMP-10	0.001	VE-Cadherin	0.04
E-Cadherin	0	IL-1 R4/ST2	0.006	MMP-7	0.006	VEGF-C	0
ErbB2	0.004	IL-10	0.004	MMP-9	0.006		
Fas/TNFRSF6	0.001	IL-13	0.012	NGF R	0.001		
Fcr RIIB/C	0	IL-13 R alpha1	0.002	Nidogen-1	0		
Ferritin	0.011	IL-15	0.002	Oncostatin M	0.002		

**Supplementary Table S4. A paired t-test ( $p < 0.05$ ) screening of 67 differentially expressed proteins between HCC tissues and para-cancer tissues.**

Protein	Paracancer average	Tumor average	Paired t-test
IL-1 R4/ST2	336.0463779	16.21405293	0.010130448
FGF-4	218.2253411	10.72610617	0.00873652
TNF-alpha	285.8060271	22.10006402	0.013324349
BDNF	431.1836713	41.42291552	0.002182451
CNTF	607.550983	73.47477619	0.000499231
I-TAC	776.4091773	131.4170628	0.002546427
I-309	425.9986592	89.20665654	0.00070884
IL-2 Rapha	338.1192257	87.64886726	0.002069127
Axl	361.157938	95.08084452	0.000474427
FGF-9	424.9490706	124.2313417	0.001089075
LYVE-1	10246.19021	3030.320145	4.41758E-08
BMP-6	310.6391964	93.34137416	0.005631068
CK beta 8-1	247.5624468	74.83836633	0.01071304
Thrombopoietin	399.0638023	132.2249122	0.002556618
IFN-gamma	513.0922839	179.2616936	0.001871168
MIP-3-alpha	269.6051823	96.97236378	0.046586751
MMP-10	265.0196791	107.9712891	0.000692854
Fcr RIIB/C	749.1250958	306.7445268	6.48872E-05
M-CSF	565.9774313	241.7376153	0.022610864
MCP-4	322.4423751	138.2385031	0.034511498
RANK	244.3703889	105.6156763	0.000156861
IGFBP-3	461.9107766	216.1715687	0.000229518
VEGF-C	305.9857297	148.0722805	5.83362E-06
Osteopontin	302.619505	151.2018506	0.000678797
Fas/TNFRSF6	9651.481857	4911.800609	0.000201688
MMP-9	5700.035759	2901.863271	0.00349502
Shh N	585.9255253	298.4519793	0.000560887
GDNF	422.2124569	224.997724	0.009164448
IL-2	978.5719818	523.2021447	0.005372663
VE-Cadherin	257.2678994	140.5429311	0.04025829
GCSF	1794.817795	997.7916186	0.002043438
E-Cadherin	1255.62963	721.9167548	0.000332997
Resistin	555.8125734	322.8116698	0.00264719
HB-EGF	429.2490597	250.5054155	0.001488933
IL-13 R alpha1	580.5901513	339.6933758	0.000260814
hCGa, intact	876.8068873	518.2890513	0.000216336
CD40 Ligand	1005.531748	594.4613733	0.001940871
HGF	8355.707414	5002.170664	0.0403311
MMP-7	455.224252	273.2685353	0.009987108
Ubiquitin+1	331.1734193	201.4129275	0.001562813
IL-13	719.5378262	442.9891118	0.019008772
IL-15	871.6415525	547.4064262	0.004317489
ErbB2	658.6006479	414.3644436	0.002931434
IL-17R	777.9376056	493.1382936	0.015663125
Thyroglobulin	436.5635659	278.1804488	0.00027515
GM-CSF	1206.000712	770.7310689	0.001549253
IL-7	1111.538437	750.5630412	0.024138875
IL-17B	391.464719	270.8815847	0.013670341
TGF-beta 1	1118.413478	775.6240628	0.008405324
IL-10	931.781696	647.6302292	0.006498194
CD23	357.5499215	250.747118	0.006356876
DKK-3	470.651783	344.0706509	0.0237993
NGF R	630.261306	471.2715402	0.000788183
Angiopoietin-1	198.5782363	148.7715895	0.017357543
TIM-1	86.78018813	382.3287542	0.009079839
EpCAM	529.6611018	2291.461443	0.023592508
GDF-15	1509.1155	5463.53949	0.001945907
ErbB3	205.9762295	687.5797862	0.019713956
CEACAM-1	1031.037409	2809.533226	0.001192134
Follistatin	1873.946527	4645.614154	0.048619006
PAI-I	482.5540539	1145.362537	0.012017737
Acrp30	24506.68502	45559.3495	0.000725954
BCAM	728.8332415	1329.85711	0.005938727
Nidogen-1	6522.389135	10766.05042	1.08749E-05
TIMP-1	828.0197193	1359.009075	0.030160722
Ferritin	326.7858257	499.2970531	0.006684934
Angiogenin	31831.61648	41646.37015	0.042907819

**Supplementary Table S5. Differentially expressed proteins of the intersection of 33 SAM ( $|\text{d-score}| > 3$ ), Wilcoxon and t-test results.**

Protein	Gene	Wilcoxon	paired t	$ \text{d-score}  > 3$	Protein	Gene	Wilcoxon	paired t	$ \text{d-score}  > 3$
Acrp30	ADIPOQ	0.001	0.000725954	4.040771312	IGFBP-3	IGFBP3	0	0.000229518	-4.025513175
Axl	AXL	0.003	0.000474427	-3.33310897	IL-13 R alpha1	IL13RA1	0.002	0.000260814	-3.419419264
CD40 Ligand	CD40LG	0.001	0.001940871	-3.900368469	IL-17R	IL17RA	0	0.015663125	-4.161035342
CEACAM-1	CEACAM1	0.001	0.001192134	3.456695869	I-TAC	CXCL11	0	0.002546427	-3.139021677
CNTF	CNTF	0.001	0.000499231	-3.520364411	LYVE-1	LYVE1	0	4.41758E-08	-7.524597203
E-Cadherin	CDH1	0	0.000332997	-3.441371544	MMP-10	MMP10	0.001	0.000692854	-3.321856001
ErbB2	ERBB2	0.004	0.002931434	-3.07203438	MMP-9	MMP9	0.006	0.00349502	-3.096942344
Fas/TNFRSF6	FAS	0.001	0.000201688	-4.289788962	NGF R	NGFR	0.001	0.000788183	-3.290269729
Fcr RIIB/C	FCGR2B	0	6.48872E-05	-4.485951627	Nidogen-1	NID1	0	1.08749E-05	5.260659198
FGF-9	FGF9	0.002	0.001089075	-3.555722531	Osteopontin	SPP1	0.001	0.000678797	-3.309735498
GCSF	CSF3	0.002	0.002043438	-3.152900461	RANK	TNFRSF11A	0.001	0.000156861	-3.660103122
GDF-15	GDF15	0.001	0.001945907	3.360832432	Resistin	RETN	0.005	0.00264719	-3.049152962
GM-CSF	CSF2RB	0	0.001549253	-3.378171995	Shh N	SHH	0.001	0.000560887	-3.793116139
HB-EGF	HBEGF	0.007	0.001488933	-3.119954508	Thyroglobulin	TG	0	0.00027515	-3.380140756
hCGa, intact	CGA	0	0.000216336	-3.107843916	Ubiquitin+1	UBA52	0.001	0.001562813	-3.132476113
I-309	CCL1	0.001	0.00070884	-3.422519102	VEGF-C	VEGFC	0	5.83362E-06	-4.613277491
IFN-gamma	IFNG	0.005	0.001871168	-3.032543676					