International Journal of General Medicine

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

Role of *CD5L* and *SRD5A2* as Prognostic Biomarkers for Hepatocellular Carcinoma

Yunxiu Luo[1,*]
Xiaopeng Huang[2,*]
Jiabin Zhan[3]
Shuai Zhang[2]

[1]Department of Radiation Oncology, Hainan Cancer Hospital, Affiliated Cancer Hospital of Hainan Medical University, Haikou, 570311, Hainan Province, People's Republic of China; [2]Department of Radiation Oncology, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, 570311, Hainan Province, People's Republic of China; [3]Department of Otolaryngology, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, 570311, Hainan Province, People's Republic of China

*These authors contributed equally to this work

Purpose: Due to the limitations of currently available biomarkers, new biomarkers are needed to accurately predict the prognosis of patients with hepatocellular carcinoma (HCC) patients.

Methods: In this study, we screened for differentially expressed genes (DEGs) in the tumor and the adjacent tissues using the four gene expression array (*GSE14520, GSE45267, GSE121248, GSE62232*) of the Gene Express Omnibus (GEO) database.

Results: Subsequently, 47 overlapping DEGs were identified in four GEO datasets, which were mostly located on chromosomes 5q and 6q, distributed in the liver and CD105-positive endothelial cells, and closely related to HCC. Function enrichment revealed 47 DEGs were related to HCC, and involved in steroid /lipid /retinol metabolism, bile secretion and p53 signalling pathway. The Kaplan–Meier plotter analysis (http://www.kmplot.com/) identified 26 and 40 genes associated with the 5-year overall survival (OS) and relapse-free survival (RFS). We found that *CD5L* and *SRD5A2* were independent prognostic factors for 5-year OS (P=0.036) and RFS (P=0.044) in HCC patients from *GSE14520*, respectively. Clinicopathological features including BCLC stage, cirrhosis, and risk signature for predicted metastasis were used to construct and validate a nomogram for 5-year OS with C-index of 0.732 and 0.717 in the training and validation cohort, respectively. *SRD5A2*, BCLC stage and gender was independent prognostic factors for RFS which were used to build a nomogram with the C-index of 0.666 and 0.682 in the training and validation cohort, respectively.

Conclusion: *CD5L* can facilitate individualized, targeted therapy for HCC patients.

Keywords: hepatocellular carcinoma, *CD5L*, prognostic, nomogram, *SRD5A2*

Introduction

Hepatocellular carcinoma (HCC) is the sixth most common malignant disease and the second leading cause of cancer-related mortality in males.[1] Despite the improvements in the clinical diagnosis and treatment strategies, the 5-year overall survival (OS) rate of HCC patients remains less than 40%.[1,2] Intrahepatic recurrence and extrahepatic metastasis are the major causes of death.[2] The identification of specific biomarkers and responsive therapeutic targets could significantly help to predict early recurrence and subsequently improve the prognosis of HCC patients.[3] High throughput gene expression profiling allows identification of potential biomarkers for constructing prognosis model and prediction of latent molecular mechanisms. In the last decade, there have been few reports of the combined use of multiple genes for prognostication of HCC.[3–5] However, there is a clear need to identify a more sensitive and specific prognostic biomarker for clinical practice.

Correspondence: Jiabin Zhan
Department of Otolaryngology, Hainan General Hospital, Hainan Affiliated Hospital of Hainan Medical University, Haikou, 570311, Hainan Province, People's Republic of China
Tel +86-18907669339
Email jiabin2021hainan@163.com

Shuai Zhang
Tel +86-1387642896
Email 46370976@qq.com

Received: 20 September 2021
Accepted: 18 November 2021
Published: 1 December 2021

CD5L is highly expressed in macrophages and acts as regulator of lipid synthesis, immune homeostasis and inflammatory response.[6–10] *CD5L* could inhibit chronic liver injury by controlling *SMAD7* expression and *TGFβ* signal pathway. But, representative analysis found *CD5L* protein had higher expression in HCC tissues than that in adjacent liver tissues, and the patients with high expression of *CD5L* had terrible prognosis comparing to those with under expression.[11] In vitro study indicated *CD5L* prompted the proliferation and colony formation of HCC cells and helped them to escape from cisplatin induced apoptosis.[11] The aim of the current study was to investigate the expression and clinical significance of *CD5L* and other differentially expressed genes (DEGs) in HCC using public databases.

Materials and Methods

Patients and Gene Expression Data

Gene expression and clinical data were obtained from the gene expression omnibus (GEO) database of the National Center for Biotechnology Information (NCBI) (accession numbers *GSE45267*, [12] *GSE62232*,[4] *GSE121248*,[13] and *GSE14520-GPL571*).[14] The data on the effect of gene expression on prognosis in liver cancer were acquired from the Kaplan-Meier plotter (KM plotter, http://www.kmplot.com/). Kaplan-Meier plotter is a public database containing 54,675 genes of 18,674 cancer patients, including 364 liver cancer cases with relapse-free and overall survival data derived from TCGA database.

Analysis of Differentially Expressed Genes (DEGs)

DEGs in HCC and precancerous samples were screened using the GEO2R online tool (https://www.ncbi.nlm.nih.gov/geo/geo2r/) through which we compared two or more groups of samples in a GEO dataset in order to identify genes that are differentially expressed across experimental conditions. The GEO2R online analysis tool was used to screen the DEGs in four datasets which detected the expression profile of paired samples from HCC and adjacent non-cancerous tissues. To take care of false-positive results, an adjusted p-value was calculated using the Benjamini and Hochberg false discovery rate. An adjusted *P*-value *<0.01, P*-value *<0.01*, and a log fold change (log FC) ≥2 were considered as the thresholds for DEGs screening. Gene with more than one probe set was average. The common DEGs among the four datasets were analyzed using a Venn diagram. The data processing has been shown in Figure S1.

Gene Function Annotation and PPI Network Assay

The database for annotation, visualization and integrated discovery (DAVID) (https://david.ncifcrf.gov/), Omicsbean (http://www.omicsbean.cn/) and KEGG Orthology-Based Annotation System (KOBAS) (http://kobas.cbi.pku.edu.cn/kobas3/) were used to determine the function enrichment for biological process (BP), molecular function (MF), and cellular component (CC), and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway of DEGs. *P<0.05* was considered to be statistically significant. The protein-protein interaction network analysis of DEGs was conducted using the String online database (http://www.string-db.org/, version 11.1). A combined score of more than 0.4 was set as the cut-off value.

Screening for Survival-Related DEGs

The univariate survival analysis of overlapped DEGs in four datasets was performed according to the K-M method using the KM plotter (http://kmplot.com/). The screening filters for gene selection were as follows: multiple genes (gene symbols), split patients by median value, threshold overall survival (OS) of 60 months, median survival, censored at threshold, the pathology and clinical features of the patients were set as default all (stage, grade, AJCC T, vascular invasion, gender, race, sorafenib treatment, alcohol consumption, hepatitis virus were set all). $P< 0.05$ was used for the purpose of analyzing significantly different. The data processing has been shown in Figure S1.

Statistical Analysis

Statistical analysis was performed using the EmpowerStats, Graph Pad Prism 7.0, IBM SPSS 23.0, and R 3.5.3 software packages. Continuous variables were assayed by two-sided Student's *t*-test, or one-way analysis of variance (ANOVA), while nonparametric or χ^2 test (Fisher exact probability method) were used for categorical variables. Normally distributed data are presented as means ± standard deviation, and the data with skewed distribution are presented as median. Factors that appeared statistically significant on univariate analysis were subjected to multivariate Cox regression analysis. *P <0.05* was considered to be statistically significant.

Figure 1 Venn diagram analysis of DEGs in different GEO datasets. Individual studies are indicated in different colors. The overlapping parts indicate DEGs common to different GSE datasets. Together, 47 DEGs were screened from four studies.
Abbreviation: DEGs, differentially expressed genes.

The R 3.5.3 software was used to construct the nomogram and the calibration curve.

Results

Identification of DEGs with GEO2R and Analysis of Genetic Characteristics

On analysis, 241, 346, 200 and 124 DEGs were identified in *GSE14520, GSE45267, GSE62232*, and *GSE121248* datasets, respectively (Figure 1). The four datasets contained 47 overlapped DEGs. Out of these, 81% of DEGs (38/47) were downregulated, and 19% of DEGs (9/47) were upregulated. The 47 DEGs were subsequently analyzed. The *TOP2A, CDKN3*, and *IGFBP3* genes were over-expressed, and *CXCR12, CXCR14, CD5L, LCAT* were under-expressed (Table 1). The analysis of genetic features of the DEGs indicated that most of them were located in the chromosome regions 5q, 6p26, 11p, and 5p (Figure 2A) and distributed in the liver and CD105-positive endothelial cells (Figure 2B). Downregulated DEGs were strongly linked to HCC (Figure 2C), while the upregulated DEGs were likely associated with the pathogenesis of non-alcoholic steatohepatitis (NASH) and breast cancer (Figure 2D).

Functional Classification of DEGs and Protein-Protein Interaction Network

Using the Gene Ontology (GO) annotation, the number of DEGs in each GO term, including CC, BP, MF and corresponding KEGG pathways were also identified. The DEGs were predominantly involved in protein binding, monooxygenase activity, oxidoreductase activity, and the reduction of molecular oxygen (Figure S2). The BP included steroid metabolism, lipid metabolism, monocarboxylic acid metabolism, and response to chemicals (Figure S2). The KEGG pathways included retinol metabolism, bile secretion, tryptophan metabolism, cytokine-cytokine receptor interaction, and the p53 signaling pathway, as shown in Table S1. To understand the interactions involving DEGs, PPI, (Figure 3), network was generated using String 11.1. The detected interactions indicated that *TOP2A* and *CDKN3* were the key nodes in the network. They participated in the regulation of cell cycle and were related to carcinogenesis. *TOP2A* was also involved in metabolism of protein and *CDKN3* was associated with HCC.[15–18]

Identification of the Prognosis-Related DEGs Through KM Plotter Analysis (TCGA Cohort)

The K-M method was used in the present data to gain insight into the association between the DEGs and the prognosis (5-year OS, RFS) of patients with HCC. A total of 47 DEGs were used to analyze their association with 5-year OS and RFS in 364 and 313 patients from TCGA database, respectively. Twenty-six genes (*BBOX1, ZG16, ASPM, CD5L, TOP2A, DTL, ESR1, FCN2, NAT2, LCAT, GLS2, GHR*, as seen in Table 2) were associated with 5-year OS, and 40 genes were

Table 1 Forty-Seven DEGs

Down-Regulated	*BBOX1*	*CYP26A1*	*CXCL14*	*CYP39A1*	*ZG16*	*SLCO1B3*	*CYP1A2*	*ACSM3*	*CD5L*	*SLC22A1*
	FCN3 *DNASE1L3* *LPA*	*GBA3* *NAT2* *GHR*	*CYP2B6* *RACGAP1* *CLEC1B*	*KMO* *CXCL12* *KCNN2*	*LCAT* *CRHBP* *GLS2*	*CLEC4M* *MT1M* *APOF*	*ESR1* *GLYAT* *ECT*	*HAMP* *SRD5A2* *MT1F*	*IL1RAP* *SRPX* *HAO2*	*FCN2*
Up-regulated	*ASPM*	*PRC1*	*RRM2*	*TOP2A*	*CAP2*	*DTL*	*GPC3*	*IGFBP3*	*CDKN3*	

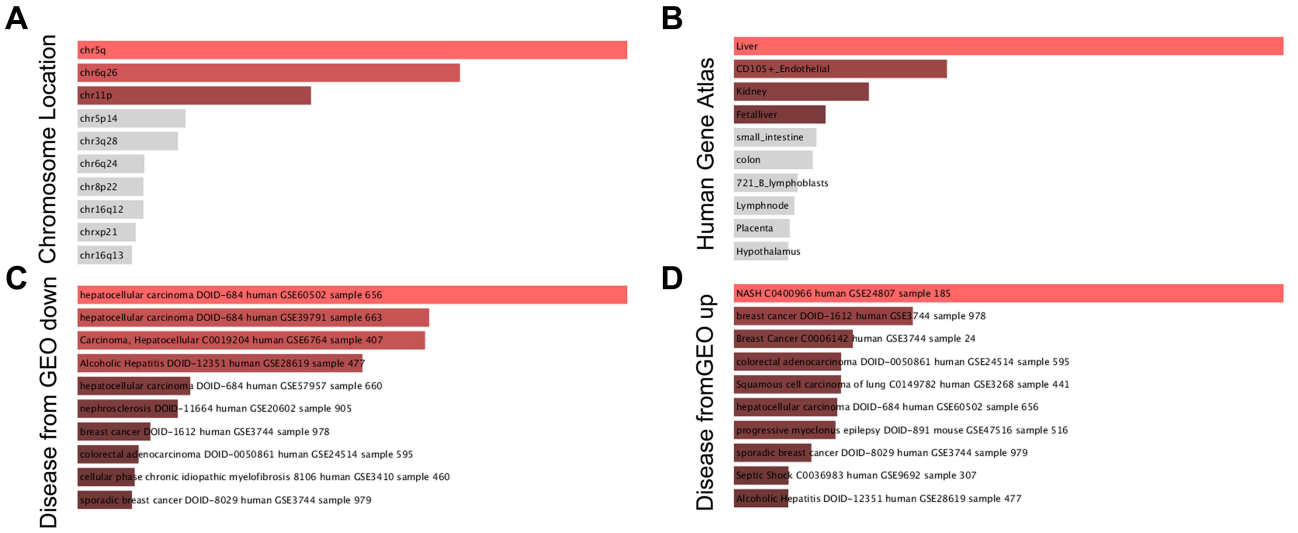

Figure 2 The GO analysis of the 47 identified DEGs performed with the enrichR software. (**A**) chromosome distribution of DEGs; (**B**) anatomic distribution of DEGs; (**C**) down-regulated expression of DEGs in HCC; (**D**) up-regulated expression of DEGs in HCC.
Abbreviation: DEG, differentially expressed gene; GO gene ontology.

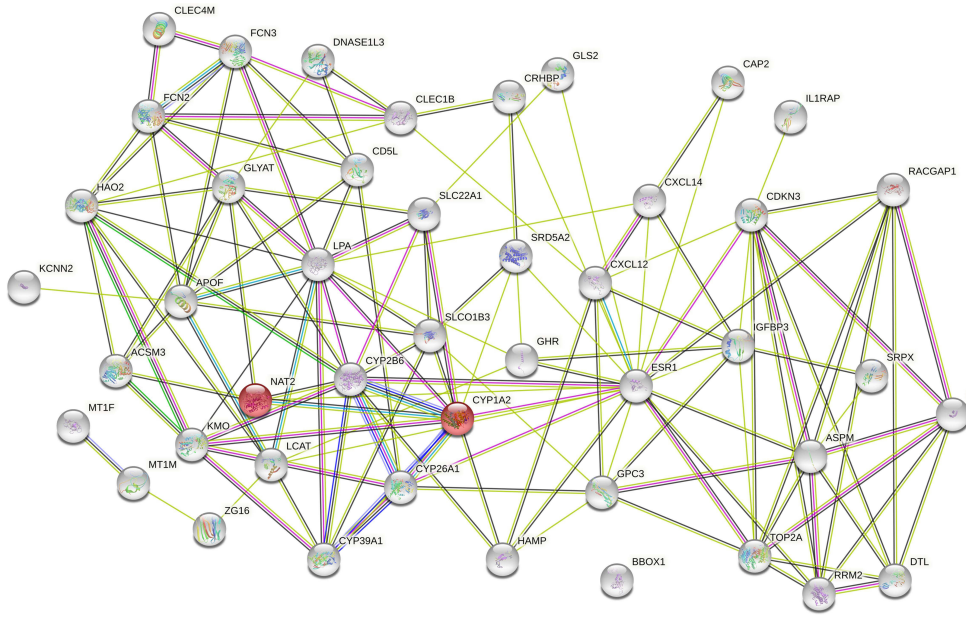

Figure 3 Analysis of PPI network of the 47 identified DEGs using Sting 11.1.
Abbreviations: DEG, differentially expressed gene; PPI, protein-protein interaction.

linked to RFS in HCC patients (Table 2). Twenty-two genes related to both OS and RFS were analyzed further using a regression model, as detailed below.

Validation of Candidate Genes and Identification of Independent Prognostic Factors (GSE14520 Cohort)

To confirm the relationship between the DEGs and prognosis, validation EGs was performed using different cases from *GSE14520* cohort. The clinicopathologic features and 22 genes linked to 5-year OS were analyzed by univariate Cox regression, and the results are reported in Table 3. Tumor stage (AJCC TNM stage, BCLC stage, CLIP stage), risk signature for predicted metastasis (RSPM), tumor size, multiple nodular, cirrhosis, *AFP, ZG16, ASPM, ACSM3, CD5L, PRC1, DNASE1L3, NAT2, RACGAP1, TOP2A, GLYAT, SRD5A2, SLC22A1, LCAT, ESR1*, and *GHR* were associated with 5-year OS.

Table 2 Validation of DEGs Correlated to RFS and 5-Year OS with K-M Plotter

RFS						5y OS					
	Median		HR	95% CI	p		Median		HR	95% CI	p
	Low Expression	High Expression					Low Expression	High Expression			
BBOX1	25.13	40.97	7.30E-01	0.52–1.03	6.80E-02	*BBOX1*	13.10	25.70	6.50E-01	0.45–0.93	1.70E-02
CYP26A1	13.27	37.67	5.20E-01	0.37–0.72	1E-04	*ZG16*	46.60	59.70	6.40E-01	0.44–0.92	1.40E-02
CXCL14	21.47	50.30	5.50E-01	0.37–0.83	3.30E-03	*ASPM*	31.00	13.10	1.70E+00	1.18–2.44	4.00E-03
CYP39A1	15.97	33.00	6.80E-01	0.49–0.94	2.00E-02	*ACSM3*	13.70	27.60	5.80E-01	0.4–0.83	3.10E-03
ZG16	19.73	34.40	6.90E-01	0.49–0.99	4.50E-02	*CD5L*	14.60	25.20	6.10E-01	0.42–0.87	6.20E-03
ASPM	33.00	13.27	1.66E+00	1.18–2.33	3.10E-03	*LPA*	12.40	27.90	5.60E-01	0.39–0.80	1.40E-03
SLCO1B3	21.93	37.23	6.80E-01	0.48–0.99	4.10E-02	*GHR*	12.40	26.70	5.10E-01	0.36–0.74	2.70E-04
CD5L	10.40	37.23	4.60E-01	0.33–0.65	7.8E-06	*CLEC1B*	18.20	25.70	6.10E-01	0.39–0.96	3.00E-02
LPA	15.17	42.63	5.30E-01	0.38–0.74	1.00E-05	*PRC1*	31.00	12.20	1.91E+00	1.32–2.75	4.20E-04
GHR	12.37	37.23	5.40E-01	0.39–0.76	3.00E-04	*DNASE1L3*	11.50	38.20	4.20E-01	0.29–0.6	1.90E-06
CLEC1B	16.60	50.30	4.70E-01	0.34–0.65	3.3E-06	*NAT2*	14.00	26.70	6.60E-01	0.46–0.95	2.30E-02
PRC1	36.10	12.87	1.81E+00	1.29–2.54	5.00E-04	*RACGAP1*	31.00	12.20	1.92E+00	1.33–2.77	4.10E-04
DNASE1L3	10.03	42.63	4.00E-01	0.28–0.55	1.6E-08	*RRM2*	27.90	12.70	1.70E+00	1.18–2.44	3.80E-03
RACGAP1	37.23	11.97	2.00E+00	1.41–2.83	7.3E-05	*CRHBP*	13.80	25.70	5.90E-01	0.41–0.85	3.70E-03
CXCL12	21.20	50.30	5.30E-01	0.37–0.76	4.00E-04	*TOP2A*	31.00	11.60	1.95E+00	1.36–2.82	2.50E-04
RRM2	36.10	16.37	1.70E+00	1.22–2.37	1.00E-03	*GLYAT*	13.80	24.10	6.80E-01	0.47–0.97	3.20E-02
CRHBP	10.70	36.10	4.70E-01	0.33–0.68	3.8E-05	*SRD5A2*	14.20	27.60	6.50E-01	0.45–0.93	1.80E-02
TOP2A	36.10	11.83	1.90E+00	1.36–2.66	1.00E-04	*FCN2*	13.70	25.70	6.40E-01	0.44–0.91	1.30E-02
GLYAT	16.83	30.10	6.90E-01	0.48–1	4.00E-02	*SLC22A1*	12.40	31.00	4.80E-01	0.33–0.71	7.00E-05
SRPX	21.87	42.87	6.40E-01	0.44–0.93	2.00E-02	*DTL*	59.70	46.60	1.63E+00	1.14–2.35	7.30E-03
FCN2	21.93	34.40	6.90E-01	0.5–0.97	3.00E-02	*HAO2*	13.90	27.90	6.20E-01	0.43–0.89	8.80E-03
SLC22A1	15.07	42.63	5.10E-01	0.36–0.7	3.9E-05	*FCN3*	13.10	25.70	5.80E-01	0.41–0.84	3.20E-03
CAP2	37.67	21.30	1.42E+00	1.02–1.96	4.00E-02	*GBA3*	13.80	25.60	6.30E-01	0.44–0.91	1.20E-02
DTL	37.67	15.07	1.90E+00	1.35–2.66	2.00E-05	*LCAT*	13.70	31.00	4.90E-01	0.34–0.71	1.10E-04
HAO2	21.23	40.97	5.90E-01	0.43–0.82	2.00E-04	*ESR1*	12.70	38.30	4.60E-01	0.31–0.66	2.40E-05
KCNN2	21.23	37.67	7.10E-01	0.51–0.99	4.00E-02	*GLS2*	13.80	25.60	6.60E-01	0.46–0.95	2.40E-02
FCN3	15.97	42.87	5.30E-01	0.38–0.73	1.00E-04						
GBA3	19.73	37.67	6.60E-01	0.47–0.91	1.00E-02						
CYP2B6	15.07	30.40	6.70E-01	0.48–0.95	2.00E-02						
KMO	15.63	37.67	6.10E-01	0.44–0.85	3.00E-03						
LCAT	11.83	42.63	4.00E-01	0.29–0.57	6E-08						
CLEC4M	13.33	40.97	4.50E-01	0.32–0.63	1.8E-06						
ESR1	11.83	30.40	5.70E-01	0.41–0.8	1.00E-03						

(*Continued*)

https://doi.org/10.2147/IJGM.S337769

Table 2 (Continued).

	RFS						5y OS					
	Median		HR	95% CI	p		Median		HR	95% CI	p	
	Low Expression	High Expression					Low Expression	High Expression				
IGFBP3	37.23	21.23	1.54E+00	1.08–2.19	1.70E-02							
HAMP	21.23	40.97	6.80E-01	0.49–0.95	2.30E-02							
CDKN3	30.40	17.90	1.46E+00	1.05–2.03	2.20E-02							
GLS2	21.30	37.23	6.70E-01	0.48–0.94	1.80E-02							
APOF	11.97	36.10	5.70E-01	0.4–0.81	1.30E-03							
MT1M	21.30	36.10	7.20E-01	0.51–1.01	5.40E-02							
SRD5A2	25.13	36.10	7.10E-01	0.5–1.01	5.40E-02							

Moreover, tumor stage (AJCC TNM stage, BCLC stage), RSPM, tumor size, gender, *TOP2A, SLC22A1, LCAT, SRD5A2, CRHBP, GLYAT, NAT2, DNASE1L3, ZG16*, and *LPA* were linked to RFS.

Based on statistical significance, the BCLC stage was selected for OS and RFS analysis. The multivariate COX regression hazard analysis demonstrated that RSPM, cirrhosis, BCLC stage, and *CD5L* were independent prognostic factors for OS, while gender, BCLC stage, and *SRD5A2* were risk factors for RFS, as shown in Figure 4.

Construction and Validation of a DEGs-Based Predictive Nomogram Model Using GSE14520 Cohort

The cases of GSE14520 were divided into the training (n=144) and validation cohort (n=98). The validation cohort was used for the external validation of the nomogram model. The stage of cirrhosis (hazard ratio [HR] 4.97, 95% confidence interval [95% CI] 1.17–21.18, *P=0.028*), RSPM (HR 11.73, 95% CI 1.63–84.58, P=0.015), BCLC (HR 0.55, 95% CI 0.27–1.13, *P<0.001*), and *CD5L* (HR 0.712748, 95% CI 0.54–0.93, *P=0.036*) were found to be independent risk factors for 5-year OS (Figure 4A). BCLC stage (HR 4.48, 95% CI 1.66–7.04, *P=0.005*), gender (HR 0.42, 95% CI 0.22–0.81, *P=0.031*), and *SRD5A2* (HR 0.87, 95% CI 0.77–0.97, *P=0.044*) were risk factors for RFS (Figure 4B). The independent prognostic factors were incorporated into a nomogram to estimate the OS and RFS, respectively (Figure 5A and D). The performance of the nomogram was assessed by the index of concordance (C-index). The calibration plot displays ideal consistency for 5-y OS using bootstrap sampling, with a C-index of 0.73 (95% CI 0.71–0.90) in the training cohort and 0.72 (95% CI:0.70–0.89) in the validation cohort, respectively (Figure 5B and C). The C-index for the nomogram predicting 1-yr RFS was 0.67 (95% CI 0.68–0.87) in the training cohort and 0.68 (95% CI 0.68–0.89) in the validation cohort, respectively (Figures 5E and F).

Correlations Between CD5L, SRD5A2 and Clinicopathological Features (GSE14520)

The correlation between *CD5L* and *SRD5A2* with the clinicopathological features was investigated, and the results indicated that *CD5L* was related to the risk index for predicted metastasis (RIPM, *P*=0.015), TNM stage (*P*=0.05), but not to gender (*P*=0.336), cirrhosis (*P*=0.811), multiple

Table 3 Clinicopathologic Features and Their Correlation with 5-Year OS and RFS on Univariate Analysis

Univariate							
		OS			RFS		
		N (%)	β (95% CI)	*p*-value	N	β (95% CI)	*p*-value
Median		39.695+21.987	0.913 (0.899, 0.927)	<0.001***	33.310+22.883	0.932 (0.919, 0.944)	<0.001***
Age	median	50.843+10.887	0.990 (0.972, 1.008)	0.286	50.843+10.887	0.998 (0.983, 1.013)	0.775
Status							
	alive	146 (60.331%)	1		106 (43.802%)*	1	
	dead	96 (39.669%)	9.695 (6.501, 14.458)	<0.001***	136 (56.198%)	123.407 (17.177, 886.592)	<0.001***
PRMS							
	low	121 (50.000%)	1		121 (50.000%)	1	
	high	121 (50.000%)	2.251 (1.484, 3.414)	<0.001***	121 (50.000%)	1.605 (1.144, 2.253)	<0.001***
Gender							
	male	211 (87.190%)	1		211 (87.190%)	1	
	female	31 (12.810%)	0.538 (0.261, 1.110)	0.093	31 (12.810%)	0.424 (0.223, 0.807)	0.009
HBV							
	no	1 (0.415%)	1		1 (0.415%)	1	
	yes	182 (75.519%)	1.186 (0.755, 1.863)	0.459	182 (75.519%)	0.345 (0.048, 2.488)	0.291
	NA	58 (24.066%)	0.246 (0.033, 1.836)	0.172	58 (24.066%)	0.442 (0.060, 3.242)	0.422
ALT							
	low	142 (58.678%)	1		142 (58.678%)	1	
	high	100 (41.322%)	1.155 (0.772, 1.727)	0.483	100 (41.322%)	1.381 (0.986, 1.935)	0.060
Size							
	small	153 (63.485%)	1		153 (63.485%)	1	
	large	88 (36.515%)	1.960 (1.309, 2.933)	0.001**	88 (36.515%)	1.424 (1.008, 2.012)	0.044*
Multiple nodular							
	NA	190 (78.512%)	1		190 (78.512%)	1	
	no	52 (21.488%)	1.653 (1.064, 2.569)	0.025*	52 (21.488%)	1.353 (0.913, 2.005)	0.132
Cirrhosis							
	NA	19 (7.851%)	1		19 (7.851%)	1	
	no	223 (92.149%)	5.093 (1.255, 20.671)	0.023*	223 (92.149%)	2.003 (0.936, 4.287)	0.074
TNM satge							
	I	113 (46.694%)	1		113 (46.694%)	1	
	II	78 (32.231%)	1.981 (1.531, 2.561)	0.060	78 (32.231%)	1.681 (1.135, 2.491)	0.01*
	III	51 (21.074%)	3.912 (2.381, 6.425)	<0.001***	51 (21.074%)	2.597 (1.686, 4.001)	<0.001***
BCLC stage							
	0	20 (8.264%)	1		20 (8.264%)	1	
	A	169 (69.835%)	2.068 (1.647, 2.597)	0.039*	169 (69.835%)	2.208 (0.966, 5.046)	0.045*
	B	24 (9.917%)	9.200 (2.085, 40.594)	0.003**	24 (9.917%)	3.944 (1.548, 10.048)	0.004**
	C	29 (11.983%)	16.845 (3.935, 72.113)	<0.001***	29 (11.983%)	6.230 (2.526, 15.363)	<0.001***
CLIP stage							
	0	99 (41.079%)	1		99 (41.079%)	1	
	1	94 (39.004%)	1.602 (0.954, 2.687)	0.075	94 (39.004%)	1.391 (0.941, 2.056)	0.098
	2	35 (14.523%)	3.445 (1.949, 6.090)	<0.001***	35 (14.523%)	2.004 (1.225, 3.277)	0.006**

(*Continued*)

Table 3 (Continued).

Univariate							
		OS			RFS		
		N (%)	β (95% CI)	*p*-value	N	β (95% CI)	*p*-value
	3	9 (3.734%)	5.385 (2.346, 12.361)	<0.001***	9 (3.734%)	2.070 (0.885, 4.843)	0.093
	NA	4 (1.660%)	5.453 (2.250, 12.706)	<0.001***	4 (1.660%)	7.683 (2.348, 25.140)	<0.001***
AFP							
	low	132 (54.545%)	1		132 (54.545%)	1	
	high	110 (45.455%)	1.704 (1.140, 2.547)	0.009**	110 (45.455%)	1.347 (0.962, 1.886)	0.083
BBOX1		4.322+1.470	0.920 (0.797, 1.061)	0.025*	4.322+1.470	0.925 (0.823, 1.040)	0.192
ZG16		5.297+1.166	0.713 (0.589, 0.865)	0.001**	5.297+1.166	0.775 (0.665, 0.903)	0.001**
ASPM		6.796+1.337	1.230 (1.038, 1.459)	0.017*	6.796+1.337	1.095 (0.957, 1.252)	0.186
ACSM3		4.697+1.076	0.805 (0.658, 0.985)	0.035*	4.697+1.076	0.891 (0.760, 1.045)	0.156
CD5L		4.977+0.987	0.741 (0.596, 0.922)	0.007**	4.977+0.987	0.887 (0.746, 1.055)	0.174
LPA		4.617+0.997	0.844 (0.672, 1.058)	0.141	4.617+0.997	0.820 (0.677, 0.993)	0.042*
CLEC1B		4.132+0.453	0.869 (0.556, 1.360)	0.540	4.132+0.453	0.868 (0.601, 1.254)	0.450
PRC1		6.214+1.155	1.276 (1.061, 1.534)	0.01*	6.214+1.155	1.130 (0.973, 1.312)	0.109
DNASE1L3		5.325+1.257	0.733 (0.613, 0.876)	<0.001***	5.325+1.257	0.816 (0.708, 0.939)	0.005**
NAT2		5.171+1.518	0.829 (0.715, 0.962)	0.013*	5.171+1.518	0.889 (0.792, 0.998)	0.047*
RACGAP1		6.336+0.930	1.396 (1.113, 1.750)	0.004**	6.336+0.930	1.184 (0.986, 1.421)	0.070
RRM2		7.117+1.327	1.136 (0.972, 1.328)	0.109	7.117+1.327	1.066 (0.939, 1.211)	0.321
CRHBP		3.962+0.673	0.724 (0.489, 1.072)	0.107	3.962+0.673	0.724 (0.526, 0.997)	0.048*
TOP2A		6.705+1.379	1.269 (1.090, 1.478)	0.002**	6.705+1.379	1.136 (1.005, 1.285)	0.042*
GLYAT		5.418+1.596	0.773 (0.667, 0.897)	<0.001***	5.418+1.596	0.849 (0.757, 0.951)	0.005**
SRD5A2		4.966+1.473	0.852 (0.732, 0.993)	0.04*	4.966+1.473	0.884 (0.782, 1.000)	0.050
FCN2		3.840+0.277	0.600 (0.273, 1.319)	0.203	3.840+0.277	0.701 (0.373, 1.315)	0.268
SLC22A1		6.311+2.534	0.875 (0.802, 0.954)	0.003**	6.311+2.534	0.907 (0.846, 0.972)	0.006**
LCAT		6.181+1.251	1.213 (1.034, 1.424)	0.018*	6.181+1.251	1.075 (0.939, 1.231)	0.296
ESR1		4.392+0.574	0.639 (0.448, 0.911)	0.013*	4.392+0.574	0.660 (0.490, 0.889)	0.006**
GLS2		3.895+0.222	0.393 (0.145, 1.062)	0.066	3.895+0.222	0.501 (0.226, 1.112)	0.089
GHR		5.674+1.928	0.852 (0.759, 0.956)	0.006**	5.674+1.928	0.915 (0.835, 1.002)	0.055

Notes: *p<0.05; **p<0.01; ***p<0.001

nodularity (P=0.06), ALT (P=0.296), HBV (P=0.329), tumor size (P=0.827), BCLC stage (P=0.186), CLIP stage (P=0.633), and AFP (P=0.197). On the other hand, *SRD5A2* expression was associated with RIPM (P<0.001), TNM stage (P=0.012), BCLC stage (P=0.011), CLIP stage (P<0.001), and AFP (P<0.001), but not with gender (P=0.178), cirrhosis (P=0.473), multiple nodularity (P=0.118), ALT (P=0.117), HBV (P=0.308), and tumor size (P=0.068). Lastly, all 22 genes correlated with *CD5L* and *SRD5A2* (Table 4).

Discussion

The pathogenesis of HCC is related to the dysregulation of oncogenes and tumor suppressor genes, which are also

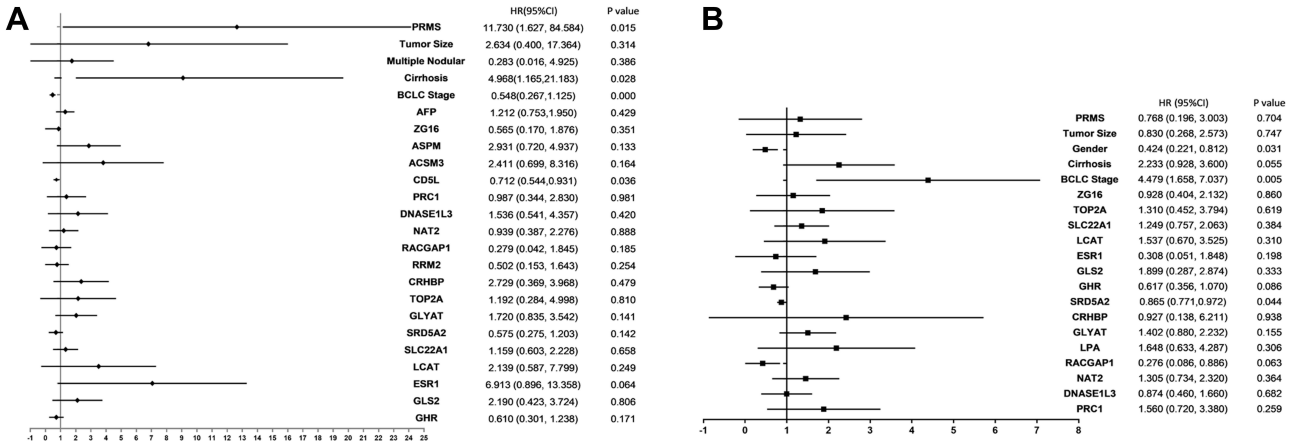

Figure 4 Forest plots summarizing the analysis of OS and RFS. Multivariate analysis of OS (**A**) and RFS (**B**) in HCC patients. The tangle squares on the transverse lines indicate the HR, and the transverse lines represent 95% CI.
Abbreviations: OS, overall survival; RFS, relapse-free survival; HCC, hepatocellular carcinoma; CI, confidence interval.

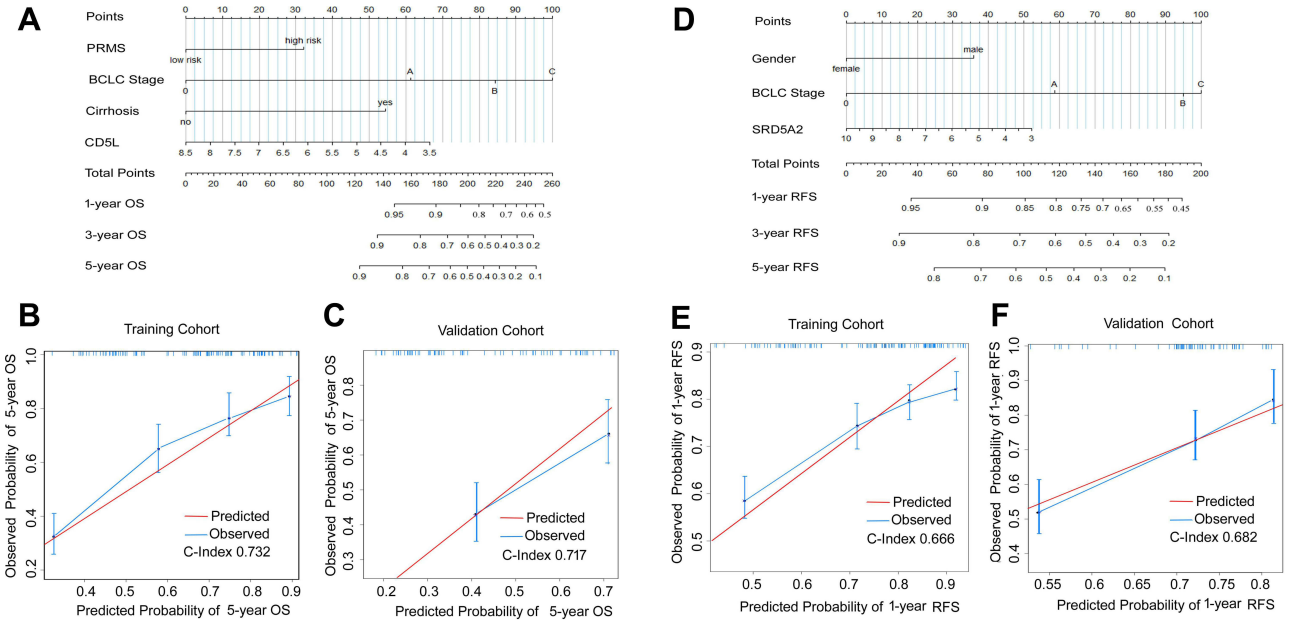

Figure 5 Nomogram to estimate the prognosis with external /internal validation. To use the nomogram, find the position of each variable on the corresponding axis and determine the number of points for each variable. The corresponding points of all the variables are added, and the probabilities of the survival are determined using the lines at the bottom of the nomogram. (**A**) nomogram for 5-year OS. (**B**) Internal calibration plot showing the performance of the proposed nomogram in predicting the 5-year OS in the training cohort (n=144). (**C**) External validation plot showing the predictive performance of the nomogram in estimating the 5-year OS in the validation cohort (n =99). The closer the blue curve is to the red line, the better is the performance. (**D**) nomogram for 1-year RFS. (**E**) the internal calibration plot shows the performance of the proposed nomogram in predicting 1-year RFS in the training cohort (n=144). (**F**) the external validation plot showing the predictive performance of the nomogram in estimating the 1-year RFS in the validation cohort (n=99).
Abbreviations: OS, overall survival; RFS, relapse-free survival.

responsible for the tumor progression and metastasis. In the current study, we identified 47 DEGs in HCC tissues, most of which were located on chromosomes 5q and 6q, and expressed mostly in the liver and CD105-positive endothelial cells. Previous studies also had similar observations with the tumor suppressor gene of non-cirrhotic HCC located on chromosome 5q and the gene related to recurrence located on chromosome 6q.[19–22] In this study, most of DEGs with low expression were located on chromosome 5q, suggesting that the decreased expression of these genes promoted hepatocarcinogenesis.[21] The upregulated genes were related to NASH and breast cancer, while the downregulated genes were closely associated with HCC. The GO and KEGG pathway analysis revealed

Table 4 Correlation Between *CD5L* and *SRD5A2* Expression and Clinicopathologic Features

		CD5L			*SRD5A2*		
		Low Expression	**High Expression**	***P*-value**	**Low Expression**	**High Expression**	***P*-value**
N		125	122		124	123	
OS	median	59.2	inf	0.004**	26.4	51.1	0.003**
RFS	median	28.7	48	0.101	28.917 ± 23.699	37.703 ± 21.239	0.013*
Age	mean±sd	50.545 ± 11.191	51.140 ± 10.613	0.672	49.521 ± 10.945	52.165 ± 10.711	0.059
Status-OS				0.018*			0.018*
	alive	64 (52.893%)	82 (67.769%)		64 (52.893%)	82 (67.769%)	
	dead	57 (47.107%)	39 (32.231%)		57 (47.107%)	39 (32.231%)	
Status-RFS				0.3			0.12
	no	49 (40.496%)	57 (47.107%)		47 (38.843%)	59 (48.760%)	
	relapse	72 (59.504%)	64 (52.893%)		74 (61.157%)	62 (51.240%)	
PRMS				0.015*			<0.001***
	low	51 (42.149%)	70 (57.851%)		40 (33.058%)	81 (66.942%)	
	high	70 (57.851%)	51 (42.149%)		81 (66.942%)	40 (33.058%)	
Gender				0.336			0.178
	male	108 (89.256%)	103 (85.124%)		102 (84.298%)	109 (90.083%)	
	female	13 (10.744%)	18 (14.876%)		19 (15.702%)	12 (9.917%)	
HBV				0.329			0.308
	no	33 (27.049%)	26 (21.667%)		34 (27.869%)	25 (20.833%)	
	yes	89 (72.951%)	94 (78.333%)		88 (72.131%)	95 (79.167%)	
ALT				0.296			0.117
	low	67 (55.372%)	75 (61.983%)		65 (53.719%)	77 (63.636%)	
	high	54 (44.628%)	46 (38.017%)		56 (46.281%)	44 (36.364%)	
Size				0.827			0.068
	small	76 (62.810%)	77 (64.167%)		70 (57.851%)	83 (69.167%)	
	large	45 (37.190%)	43 (35.833%)		51 (42.149%)	37 (30.833%)	
Multiple nodular				0.06			0.118
	0	89 (73.554%)	101 (83.471%)		90 (74.380%)	100 (82.645%)	
	no	32 (26.446%)	20 (16.529%)		31 (25.620%)	21 (17.355%)	
Cirrhosis				0.811			0.473
	0	10 (8.264%)	9 (7.438%)		11 (9.091%)	8 (6.612%)	
	no	111 (91.736%)	112 (92.562%)		110 (90.909%)	113 (93.388%)	
TNMstage				0.05			0.012*
	I	47 (38.843%)	66 (54.545%)		47 (38.843%)	66 (54.545%)	
	II	45 (37.190%)	33 (27.273%)		40 (33.058%)	38 (31.405%)	
	III	29 (23.967%)	22 (18.182%)		34 (28.099%)	17 (14.050%)	
BCLC stage				0.186			0.011*
	0	8 (6.612%)	12 (9.917%)		7 (5.785%)	13 (10.744%)	
	A	80 (66.116%)	89 (73.554%)		78 (64.463%)	91 (75.207%)	
	B	16 (13.223%)	8 (6.612%)		14 (11.570%)	10 (8.264%)	
	C	17 (14.050%)	12 (9.917%)		22 (18.182%)	7 (5.785%)	
CLIP stage				0.633			<0.001***
	0	44 (36.364%)	55 (45.455%)		31 (25.620%)	68 (56.198%)	
	1	51 (42.149%)	44 (36.364%)		56 (46.281%)	39 (32.231%)	
	2	19 (15.702%)	16 (13.223%)		23 (19.008%)	12 (9.917%)	

(*Continued*)

Table 4 (Continued).

		CD5L			*SRD5A2*		
		Low Expression	**High Expression**	***P*-value**	**Low Expression**	**High Expression**	***P*-value**
	3	7 (5.785%)	7 (4.958%)		11 (9.091%)	2 (1.653%)	
AFP	low high	61 (50.413%) 60 (49.587%)	71 (58.678%) 50 (41.322%)	0.197	43 (35.537%) 78 (64.463%)	89 (73.554%) 32 (26.446%)	<0.001***
BBOX1		4.227 ± 1.476	4.405 ± 1.441	0.34	3.786 ± 1.036	4.848 ± 1.623	<0.001***
ZG16		5.114 ± 1.106	5.471 ± 1.190	0.015*	4.873 ± 0.973	5.711 ± 1.184	<0.001***
ASPM		6.938 ± 1.140	6.651 ± 1.500	0.092	7.032 ± 1.269	6.558 ± 1.363	0.005**
ACSM3		4.462 ± 0.969	4.924 ± 1.117	<0.001***	4.237 ± 0.786	5.146 ± 1.122	<0.001***
CD5L		4.208 ± 0.340	5.739 ± 0.806	<0.001***	4.746 ± 0.895	5.185 ± 1.022	<0.001***
LPA		4.472 ± 0.950	4.762 ± 1.026	0.022**	4.373 ± 0.839	4.859 ± 1.084	<0.001***
CLEC1B		4.008 ± 0.258	4.254 ± 0.558	<0.001***	4.013 ± 0.295	4.247 ± 0.540	<0.001***
PPRC1		6.606 ± 0.999	5.820 ± 1.177	<0.001***	6.480 ± 1.020	5.954 ± 1.229	<0.001***
DNASE1L3		4.713 ± 0.948	5.929 ± 1.234	<0.001***	4.825 ± 1.075	5.806 ± 1.233	<0.001***
NAT2		4.742 ± 1.103	5.622 ± 1.739	<0.001***	4.452 ± 1.051	5.908 ± 1.561	<0.001***
RACGAP1		6.539 ± 0.851	6.135 ± 0.966	<0.001***	6.581 ± 0.861	6.095 ± 0.937	<0.001***
RRM2		7.424 ± 1.188	6.818 ± 1.392	<0.001***	7.275 ± 1.244	6.973 ± 1.392	0.073
CRHBP		3.741 ± 0.325	4.180 ± 0.836	<0.001***	3.781 ± 0.383	4.135 ± 0.828	<0.001***
TOP2A		7.087 ± 1.179	6.323 ± 1.461	<0.001***	7.117 ± 1.230	6.298 ± 1.401	<0.001***
GLYAT		5.205 ± 1.562	5.632 ± 1.589	0.034*	4.152 ± 0.325	6.690 ± 1.307	<0.001***
SRD5A2		4.613 ± 1.207	5.325 ± 1.631	<0.001***	4.560 ± 1.204	5.372 ± 1.607	<0.001***
FCN2		3.773 ± 0.187	3.905 ± 0.333	<0.001***	3.768 ± 0.197	3.909 ± 0.324	<0.001***
SLC22A1		5.766 ± 2.278	6.849 ± 2.664	<0.001***	5.010 ± 1.927	7.602 ± 2.400	<0.001***
LCAT		6.546 ± 1.182	5.831 ± 1.201	<0.001***	6.440 ± 1.122	5.944 ± 1.310	0.002**
ESR1		4.205 ± 0.474	4.583 ± 0.602	<0.001***	4.239 ± 0.506	4.546 ± 0.595	<0.001***
GLS2		3.849 ± 0.213	3.939 ± 0.221	0.001**	3.814 ± 0.173	3.973 ± 0.235	<0.001***
GHR		5.451 ± 1.903	5.914 ± 1.944	0.06	5.052 ± 1.601	6.312 ± 2.036	<0.001***

Notes: *p<0.05; **p<0.01; ***p<0.001.

that the 47 aberrant genes were involved in pathways of steroid, lipid, and retinol metabolism, bile secretion, cytokine-cytokine interaction, and p53 signalling, which are implicated in the cell cycle progression, proliferation, and invasion of cancer cells. The above biological phenotypes were in conformity to the role of the hub genes (*TOP2A, ESR1, CDKN3, and PRC1*).

In this study, 26 and 40 genes related to 5-year OS and RFS in HCC were identified using K-M plotter through the third-party database (TCGA cohort database), respectively. Subsequent analysis demonstrated that *CD5L* and *SRD5A2* were independent risk factors for 5-year OS and RFS, respectively. Importantly, all independent prognostic factors were combined to build a predictive nomogram

model, which was effective in predicting the prognosis of HCC patients.

CD5L is mostly expressed in the macrophages, lymphoid and inflamed tissues and regulates inflammatory responses and lipid synthesis.[11,23,26] It is also known as the inhibitor of apoptosis in macrophages. It promotes macrophage survival by protecting them from the apoptotic effects of oxidized lipids in atherosclerosis.[10] Moreover, *CD5L* is involved in the early response to the infection by bacteria and other pathogens, where it acts as a pattern recognition receptor and activates autophagy.[10,24] *CD5L* also controls the metabolic switch in T-helper Th17 cells and regulates their expression of pro-inflammatory genes.[9,23] *CD5L* accumulation on the hepatic surface could inhibit chronic liver injury by attenuating CCL4-induced injury and fibrosis, repressing TGF-ß signal and immune cell infiltration.[6] Circulating *CD5L* potentially protects from the development of fatty liver and HCC.[23] The current assay also identified *CD5L* was associated with improved OS in HCC. Contrastingly, Aran et al observed that *CD5L* was upregulated in HCC and it enhanced HCC cell growth and antiapoptotic responses by binding to HSPA5 (GRP78).[11] It was also argued that *CD5L* was more suitable for HCC or cirrhosis accompanied by a viral infection than in the absence of an inflammatory response because *CD5L* itself was a factor associated with immune regulation.[9,23]

Previous studies pointed to the *CD5L*-related modulation of immune responses in malignancies, such as lung adenocarcinoma and HCC. Interestingly, *CD5L* has opposite effects on these tumors, promoting lung cancer but inhibiting liver cancer. It was established that *CD5L* accumulates on the surface of transformed hepatocytes and induces necrosis of the tumor. *CD5L*-deficient mice were susceptible to HCC and formed multiple liver tumors after feeding with a high-fat diet for one year. It was found that mouse *CD5L* was internalized together with CD36 by normal hepatocytes and modulated intracellular lipid metabolism.[24] Therefore, *CD5L* may serve as a potential target for the treatment of HCC. In humans, *CD5L* protein is present at a high concentration in the serum, especially in women. However, *CD5L* peaks in women in their 20s and decreases with age. Several proteomic assays pointed to the *CD5L* protein as a putative biomarker for inflammatory conditions as well as liver diseases.[25] The continuous inflammation caused by liver damage due to hepatitis virus infection, alcohol abuse, and NASH leads to hepatic fibrosis, which frequently triggers cirrhosis and, ultimately leads to HCC. Based on these findings, hCD5L can also be considered as a plasma biomarker for early detection of liver fibrosis and HCC, and the ratio of hCD5L-to- liver marker score might discriminate between HCC and non-HCC patients. HCC. Thus, *CD5L* may serve as a potential target in the development of HCC treatment. Future studies are required to verify the role of *CD5L* in HCC and develop strategies to alter its expression.

SRD5A2 is expressed in androgen-dependent tissues and responsible for converting testosterone to the more metabolically active dihydrotestosterone. Therefore, *SRD5A2 V89L* gene polymorphism has been associated with breast cancer and prostate cancer in previous studies.[26–28] In the present study, we found that *SRD5A2* was an independent risk factor for RFS in HCC. Liver regulates the metabolism and activity of sex hormones. Hence, *SRD5A2* might be aberrantly expressed in HCC and serve as an useful biomarker for early diagnosis of HCC.[29–31] However, in the current study, the C-index of *SRD5A2* was 0.67 and 0.68 indicating that *SRD5A2* may not have an adequate performance for the prognostic predictive nomogram. Future studies are required to delineate the role of *SRD5A2* in predicting the recurrence and survival of HCC patients.

Conclusion

In this study, 47 genes associated with HCC were identified, with most of them being located on chromosomes 5q and 6q. The potential pathways involving these genes were steroid metabolism, lipid metabolism, retinol metabolism, bile secretion, and p53 signalling pathway. There were 26 and 40 genes associated with the 5-year OS and RFS of HCC patients, respectively. Among them, *CD5L* and *SRD5A2* were independent risk factors for 5-year OS and RFS. A nomogram model combining *CD5L*, cirrhosis, RSPM and BCLC stage was constructed for accurate prognostication of patients with HCC. *CD5L* might be useful as a potential biomarker for HCC.

Abbreviations

AFP, Alpha-feto protein; AJCC, American Joint Committee on Cancer; ALT, Alanine aminotransferase; AST, Aspartate aminotransferase; BCLC, Barcelona Clinic Liver Cancer; CLIP, Cancer of the Liver Italian Program; DEG, differentially expressed gene; GEO, Gene Express Omnibus; GO, Gene Ontology; HCC, hepatocellular carcinoma; KEGG, Kyoto Encyclopedia of

Genes and Genomes; NASH, Non-alcoholic steatohepatitis; OS, overall survival; RFS, relapse-free survival; RSPM, risk signature for predicted metastasis; CD5L, CD5 Antigen-Like; SMAD7, SMAD Family Member 7; TGFβ, Transforming Growth Factor Beta; TCGA, The Cancer Genome Atlas; TOP2A, DNA Topoisomerase II Alpha; CDKN3, Cyclin-Dependent Kinase Inhibitor 3; IGFBP3, Insulin-Like Growth Factor Binding Protein 3; CXCR12, C-X-C Motif Chemokine Ligand 12; CXCL14, C-X-C Motif Chemokine Ligand 14; LCAT, Lecithin-Cholesterol Acyltransferase; BBOX1, Gamma-Butyrobetaine Hydroxylase 1.

Disclosure

Yunxiu Luo and Xiaopeng Huang are co-first authors for this study. The authors have no conflicts of interest in this work.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. doi:10.3322/caac.21492
2. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. *Lancet*. 2018;391(10127):1301–1314. doi:10.1016/S0140-6736(18)30010-2
3. Takeda H, Nishikawa H, Osaki Y. The new era of precision medicine in hepatocellular carcinoma: the urgent need for promising biomarkers. *Hepatobiliary Surg Nutr*. 2018;7(6):490–491. doi:10.21037/hbsn.2018.08.06
4. Schulze K, Imbeaud S, Letouzé E, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet*. 2015;47(5):505–511. doi:10.1038/ng.3252
5. Grinchuk OV, Yenamandra SP, Iyer R, et al. Tumor-adjacent tissue co-expression profile analysis reveals pro-oncogenic ribosomal gene signature for prognosis of resectable hepatocellular carcinoma. *Mol Oncol*. 2018;12(1):89–113. doi:10.1002/1878-0261.12153
6. Bárcena C, Aran G, Perea L, et al. CD5L is a pleiotropic player in liver fibrosis controlling damage, fibrosis and immune cell content. *EBioMedicine*. 2019;43:513–524. doi:10.1016/j.ebiom.2019.04.052
7. Gebe JA, Kiener PA, Ring HZ, Li X, Francke U, Aruffo A. Molecular cloning, mapping to human chromosome 1 q21-q23, and cell binding characteristics of Spalpha, a new member of the scavenger receptor cysteine-rich (SRCR) family of proteins. *J Biol Chem*. 1997;272(10):6151–6158. doi:10.1074/jbc.272.10.6151
8. Sanchez-Moral L, Ràfols N, Martori C, Paul T, Téllez É, Sarrias M-R. Multifaceted roles of CD5L in infectious and sterile inflammation. *Int J Mol Sci*. 2021;22(8):4076. doi:10.3390/ijms22084076
9. Sanjurjo L, Aran G, Roher N, Valledor AF, Sarrias M-R. AIM/CD5L: a key protein in the control of immune homeostasis and inflammatory disease. *J Leukoc Biol*. 2015;98(2):173–184. doi:10.1189/jlb.3RU0215-074R
10. Sanjurjo L, Aran G, Téllez É, et al. CD5L promotes M2 macrophage polarization through autophagy-mediated upregulation of ID3. *Front Immunol*. 2018;9:480. doi:10.3389/fimmu.2018.00480
11. Aran G, Sanjurjo L, Bárcena C, et al. CD5L is upregulated in hepatocellular carcinoma and promotes liver cancer cell proliferation and antiapoptotic responses by binding to HSPA5 (GRP78). *FASEB J*. 2018;32(7):3878–3891. doi:10.1096/fj.201700941RR
12. Wang H-W, Hsieh T-H, Huang S-Y, et al. Forfeited hepatogenesis program and increased embryonic stem cell traits in young hepatocellular carcinoma (HCC) comparing to elderly HCC. *BMC Genomics*. 2013;14(1):736. doi:10.1186/1471-2164-14-736
13. Wang SM, Ooi LLPJ, Hui KM. Identification and validation of a novel gene signature associated with the recurrence of human hepatocellular carcinoma. *Clin Cancer Res*. 2007;13(21):6275–6283. doi:10.1158/1078-0432.CCR-06-2236
14. Roessler S, Jia H-L, Budhu A, et al. A unique metastasis gene signature enables prediction of tumor relapse in early-stage hepatocellular carcinoma patients. *Cancer Res*. 2010;70(24):10202–10212. doi:10.1158/0008-5472.CAN-10-2607
15. Chromosomal mapping of the genes for the human cell cycle proteins cyclin C (CCNC), cyclin E (CCNE), p21 (CDKN1) and KAP (CDKN3) - PubMed [Internet]; [cited August 5 2020]. Available from: https://pubmed.ncbi.nlm.nih.gov/7698009/. Accessed November 22, 2021.
16. Wang L, Sun L, Huang J, Jiang M. Cyclin-dependent kinase inhibitor 3 (CDKN3) novel cell cycle computational network between human non-malignancy associated hepatitis/cirrhosis and hepatocellular carcinoma (HCC) transformation. *Cell Prolif*. 2011;44(3):291–299. doi:10.1111/j.1365-2184.2011.00752.x
17. Labbé DP, Sweeney CJ, Brown M, et al. TOP2A and EZH2 provide early detection of an aggressive prostate cancer subgroup. *Clin Cancer Res*. 2017;23(22):7072–7083. doi:10.1158/1078-0432.CCR-17-0413
18. Wong N, Yeo W, Wong W-L, et al. TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance. *Int J Cancer*. 2009;124(3):644–652. doi:10.1002/ijc.23968
19. Schaefer I-M, Schweyer S, Kuhlgatz J. Chromosomal imbalances in primary hepatic carcinosarcoma. *Hum Pathol*. 2012;43(8):1328–1333. doi:10.1016/j.humpath.2011.11.007
20. Ding SF, Habib NA, Dooley J, Wood C, Bowles L, Delhanty JD. Loss of constitutional heterozygosity on chromosome 5q in hepatocellular carcinoma without cirrhosis. *Br J Cancer*. 1991;64(6):1083–1087. doi:10.1038/bjc.1991.468
21. Raidl M, Pirker C, Schulte-Hermann R, et al. Multiple chromosomal abnormalities in human liver (pre)neoplasia. *J Hepatol*. 2004;40(4):660–668. doi:10.1016/j.jhep.2003.12.020
22. Nishida N, Nishimura T, Ito T, Komeda T, Fukuda Y, Nakao K. Chromosomal instability and human hepatocarcinogenesis. *Histol Histopathol*. 2003;18(3):897–909. doi:10.14670/HH-18.897
23. Maehara N, Arai S, Mori M, et al. Circulating AIM prevents hepatocellular carcinoma through complement activation. *Cell Rep*. 2014;9:61–74.
24. Sanjurjo L, Amézaga N, Aran G, et al. The human CD5L/AIM-CD36 axis: a novel autophagy inducer in macrophages that modulates inflammatory responses. *Autophagy*. 2015;11(3):487–502. doi:10.1080/15548627.2015.1017183
25. Gray J, Chattopadhyay D, Beale GS, et al. A proteomic strategy to identify novel serum biomarkers for liver cirrhosis and hepatocellular cancer in individuals with fatty liver disease. *BMC Cancer*. 2009;9(1):271. doi:10.1186/1471-2407-9-271
26. Zhang D, Li Q, Qu H-C, Yu T, Liu Y-R. Associations between the SRD5A2 gene V89L and TA repeat polymorphisms and breast cancer risk: a meta-analysis. *Genet Mol Res*. 2015;14(3):9004–9012. doi:10.4238/2015.August.7.9

27. Dušenka R, Tomaškin R, Kliment J, et al. Polymorphism of the SRD5A2 gene and the risk of prostate cancer. *Mol Med Rep*. 2014;10(6):3151–3156. doi:10.3892/mmr.2014.2621
28. Francis A, Sarkar S, Pooja S, et al. SRD5A2 gene polymorphisms affect the risk of breast cancer. *Breast*. 2014;23(2):137–141. doi:10.1016/j.breast.2013.11.010
29. Moribe T, Iizuka N, Miura T, et al. Identification of novel aberrant methylation of BASP1 and SRD5A2 for early diagnosis of hepatocellular carcinoma by genome-wide search. *Int J Oncol*. 2008;33(5):949–958.
30. Iizuka N, Oka M, Sakaida I, et al. Efficient detection of hepatocellular carcinoma by a hybrid blood test of epigenetic and classical protein markers. *Clin Chim Acta*. 2011;412(1–2):152–158. doi:10.1016/j.cca.2010.09.028
31. Tsunedomi R, Ogawa Y, Iizuka N, et al. The assessment of methylated BASP1 and SRD5A2 levels in the detection of early hepatocellular carcinoma. *Int J Oncol*. 2010;36(1):205–212.