Screening Hub Genes as Prognostic Biomarkers of Hepatocellular Carcinoma by Bioinformatics Analysis

Zengyuan Zhou[1],[*] ⓘ, Yuzheng Li[1],[*], Haiyue Hao[2],
Yuanyuan Wang[1], Zihao Zhou[1], Zhipeng Wang[2], and Xia Chu[1]

Abstract

Hepatocellular carcinoma (HCC) is a widespread, common type of cancer in Asian countries, and the need for biomarker-matched molecularly targeted therapy for HCC has been increasingly recognized. However, the effective treatment for HCC is unclear. Therefore, identifying additional hub genes and pathways as novel prognostic biomarkers for HCC is necessary. In this study, the expression profiles of GSE121248, GSE45267 and GSE84402 were obtained from the Gene Expression Omnibus (GEO), including 132 HCC and 90 noncancerous liver tissues. Differentially expressed genes (DEGs) between HCC and noncancerous samples were identified by GEO2 R and Venn diagrams. In total, 109 DEGs were identified in these datasets, including 24 upregulated genes and 85 downregulated genes. Subsequently, Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) preliminary analyses of the DEGs were performed using DAVID. The protein–protein interaction (PPI) network of the DEGs was constructed with the Search Tool for the Retrieval of Interacting Genes (STRING) and visualized in Cytoscape. Module analysis of the PPI network was performed using MCODE to get hub genes. Moreover, the influence of the hub genes on overall survival was determined with Kaplan–Meier plotter. All hub genes were analyzed by Gene Expression Profiling Interactive Analysis (GEPIA) and KEGG. Overall, the hub genes *DTL*, *CDK1*, *CCNB1*, *RACGAP1*, *ECT2*, *NEK2*, *BUB1B*, *PBK*, *TOP2A*, *ASPM*, *HMMR*, *RRM2*, *CDKN3*, *PRC1*, and *ANLN* were upregulated in HCC, and the survival rate was lower for HCC with increased expression of these hub genes. *CCNB1*, *CDK1*, and *RRM2* were enriched in the p53 signaling pathway, and *CCNB1*, *CDK1*, and *BUB1B* were enriched in the cell cycle. In brief, we screened 15 hub genes and pathways to identify potential prognostic markers for HCC treatment. However, the specific occurrence and development of HCC with expression of the hub genes should be verified in vivo and *in vitro*.

Keywords

hepatocellular carcinoma, bioinformatics analysis, hub genes, prognostic biomarkers

Introduction

Hepatocellular carcinoma (HCC) is a lethal malignancy of the liver. Similar to other cancers, HCC is associated with potentially alterable risk factors, such as excess body weight, alcohol abuse, hepatitis B virus (HBV), hepatitis C virus, nonalcohol fatty liver disease, and certain genes[1,2]. Most cases of HCC (80%) occur in Asian countries because of chronic HBV infection and exposure to aflatoxin[3]. It is estimated that by 2030 China will have the largest number of HCC patients in the world, an increase of 82% from 2016[4]. Therefore, we need to determine the mechanism of HCC to detect and treat early HCC at the molecular level.

To date, the common treatment methods for HCC include radiofrequency ablation or resection, transplantation, and radioembolization. Patients with tumors of different stages

[1] Department of Nutrition and Food Hygiene, Public Health College, Harbin Medical University, Harbin, China
[2] Department of Medical Oncology, The Fourth Affiliated Hospital of Harbin Medical University, Harbin, China
[*] Both the authors contributed equally to this article.

Corresponding Authors:
Xia Chu, Department of Nutrition and Food Hygiene, Public Health College, Harbin Medical University, 157 Baojian Road, Nangang District, Harbin, Hei Longjiang province, 150081, P. R. China.
Email: cx831128@163.com.

Zhipeng Wang, Department of Medical Oncology, The Fourth Affiliated Hospital of Harbin Medical University, 37 Yiyuan Street, Nangang District, Harbin, Hei Longjiang province, 150001, P. R. China.
Email: 15945194815@163.com

choose different treatment methods, but intrahepatic spread and recurrence rates are still very high after 5 years[5]. In 2013, Nault et al. discovered potential biomarkers, which shifted the research focus to genes related to the pathogenesis of HCC[6]. Gores proposed that HCC needed an individualized treatment scheme and the stratification of patients according to a five-gene score to adopt different molecularly matched targeted treatments[7]. Subsequently, genomic mutations, such as those in the *TERT* promoter, *TP53*, *CTNNB1*, *FGF*, and *PTEN* anomalies, which are used in guiding biomarker-matched molecularly targeted therapy of HCC, were detected in HCC, and research on these biomarkers has recently shown significant progress[8]. Therefore, we need to identify additional dysregulated genes to find new treatment strategies to improve prognosis and to realize precision medicine.

In recent years, microarrays have rapidly developed and have become the most successful tool to allow multiple combinatorial chemistry, genomics, and proteomics assays to be carried out in parallel[9]. Currently, a large number of microarray datasets have been disclosed, but the amount of data thoroughly analyzed is lacking. However, bioinformatics analysis can be used to illustrate large and complex datasets. In our study, we examined three HCC datasets in the Gene Expression Omnibus (GEO) and identified differentially expressed genes (DEGs) through a comparison of human HCC tissue with the corresponding noncancerous liver tissue, and applied bioinformatics analysis to identify hub genes and conduct a series of functional analyses.

Materials and Methods

Datasets from the GEO

Three microarray datasets from the GPL570 platform ([HG-U133_Plus_2] Affymetrix Human Genome U133 Plus 2.0 Array) were downloaded from the GEO (http://www.ncbi.nlm.nih.gov/geo/). The GSE121248 expression profile contains 70 HCC samples and 37 noncancerous samples derived from the National Cancer Centre Singapore[10]. The GSE45267 expression profile contains 48 HCC samples and 39 noncancerous samples from National Yang-Ming University in Taiwan[11]. The GSE84402 expression profile contains 14 HCC samples and 14 noncancerous samples from the Shanghai Cancer Institute in China[12].

Inclusion Criteria of the DEGs

GEO2 R is a free download system for online data analysis in the GEO; thus, the DEGs between HCC and noncancerous liver tissues in the GSE121248, GSE45267 and GSE84402 datasets could be obtained[13]. We established the following inclusion criteria for the DEGs: upregulated genes must have a log2 fold change (logFC) ≥ 2 and an adjusted p-value < 0.05, while downregulated genes must have a logFC ≤ -2 and an adjusted p-value < 0.05. Subsequently, Venn diagrams of the up- and downregulated genes were drawn for the different databases.

Functional Annotation of the DEGs

The Database for Annotation, Visualization and Integrated Discovery (DAVID, https://david.ncifcrf.gov/version 6.8) was used to perform a preliminary analysis of the obtained DEGs with systematic and comprehensive biological function notes. The Functional Annotation Tool is the core of DAVID, which includes Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses. Through GO enrichment analysis, we can roughly compare and classify DEGs to better understand their biological characteristics[14]. The KEGG helps us to study the functional interpretation of genes and genomes as a whole network[15]. In our paper, the threshold p-value < 0.05 was considered statistically significant.

Protein–Protein Interaction Network of the DEGs

To identify the hub regulatory genes and to examine the interactions between the DEGs, a protein–protein interaction network (PPI) was generated with the Search Tool for the Retrieval of Interacting Genes/Proteins (STRING, https://string-db.org/)[16]. These genes required an interaction score ≥ 0.4 and a maximum number of interactors $= 0$, and the genes were imported into Cytoscape (version 3.6.1) with the Molecular Complex Detection (MCODE) app (version 1.5.1)[17] to screen the modules of hub genes with a degree cut-off $= 2$, haircut on, node score cut-off $= 0.2$, k-core $= 2$, and max. depth $= 100$.

Validation of the Hub Genes

Kaplan–Meier plotter (http://kmplot.com/analysis) is a website that offers an online validation of survival biomarkers and analyzes the overall survival (OS) of patients with high and low expression of certain genes. In our study, hub genes were detected, and a survival curve was drawn. The log-rank p-value ($p < 0.05$) and hazard ratio with the 95% confidence interval were also calculated. Next, Gene Expression Profiling Interactive Analysis (GEPIA) (http://gepia.cancer-pku.cn/), an open source cancer big data analysis website, was used to analyze the differential expression of 369 HCC and 160 normal tissues from The Cancer Genome Atlas (TCGA) and the Genotype-Tissue Expression (GTEx) portal. All hub genes were analyzed singly, and |log2FC| $=1$ and p-value $= 0.01$ were used as cut-off criteria[18].

Results

Identification of the DEGs between HCC and Noncancerous Tissues

In our study, we chose 132 HCC and 90 noncancerous liver tissues from three datasets, GSE121248, GSE45267, and

Table 1. Identification of the Up- and Downregulated DEGs in HCC Tissues Compared with Noncancerous Liver Tissues.

DEGs	Gene names
Upregulated	CDK1, CAP2, DTL, PEG10, RACGAP1, CTHRC1, IGF2BP3, RRM2, CCNB1, TOP2A, ASPM, HMMR, CDKN3, PBK, GPC3, SULT1C2, NEK2, ANLN, ACSL4, DUXAP10, CRNDE, BUB1B, ECT2, PRC1
Downregulated	CYP4A22///CYP4A11, CYP26A1, BBOX1, CYP2A6, SERPINE1, PGLYRP2, LINC01093, CXCL14, SLC22A1, IGF1, CYP39A1, HAO2, IGHM, FAM134B, MT1F, SLC25A47, MFSD2A, ZG16, FLJ22763, HHIP, KCNN2, SLCO1B3, CYP1A2, CNDP1, BCO2, ACSM3, FCN3, GBA3, PDGFRA, TTC36, CLEC4G, C3P1, CDH19, CYP2B6, GYS2, FOLH1B, KMO, CD5 L, LPA, GHR, CLEC1B, CXCL2, FOSB, LIFR, FAM65C, CYP2C9, CLRN3, CYP2A7, LCAT, CLEC4 M, ESR1, FOS, LOC101928916///NNMT, PLAC8, ALDOB, HAMP, DNASE1L3, DCN, NAT2, BCHE, AKR1D1, TMEM27, CRHBP, THRSP, IDO2, HGFAC, IGHA2///IGHA1///IGH, C7, ADH4, GPM6A, OIT3, MT1 M, HGF, GLYAT, CYP2B7P///CYP2B6, JCHAIN, GLS2, SRD5A2, ADRA1A, EGR1, APOF, C9, SRPX, FCN2, LINC00844

DEGs: differentially expressed genes; HCC: hepatocellular carcinoma.
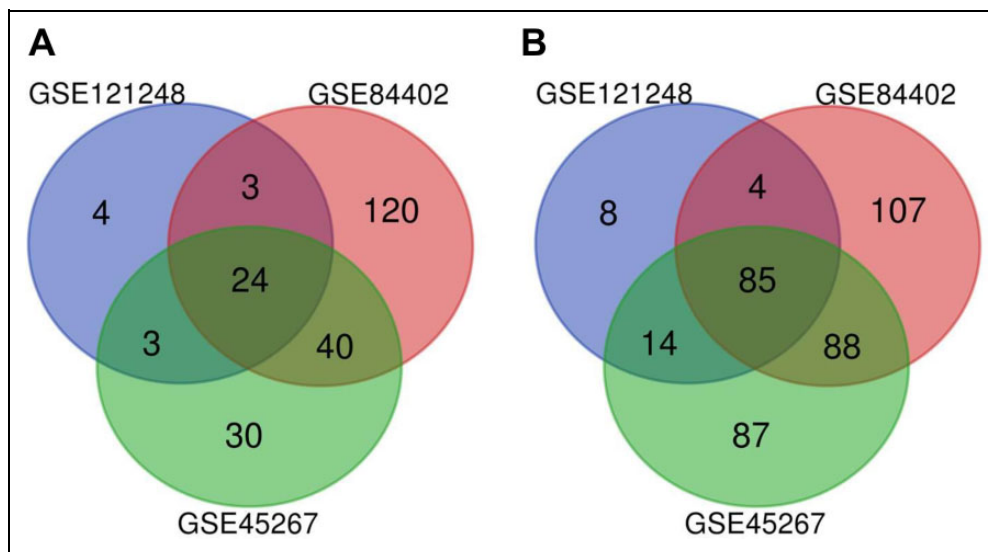

Figure 1. Identification of the DEGs between HCC and noncancerous liver tissues in three datasets via a Venn diagram. The blue circle indicates GSE121248, the green circle indicates GSE45267, and the red circle indicates GSE84402. A. Twenty-four upregulated genes had a logFC $\geq$ 2 and an adjusted p-value < 0.05. B. Eighty-five downregulated genes had a logFC $\leq$ −2 and an adjusted p-value < 0.05.

GSE84402, which were analyzed by GEO2 R and on the basis of filter criteria; 176, 480, and 585 DEGs were obtained, respectively (Additional files 1, 2, and 3). However, we found 109 DEGs with repeated emergence in these datasets, including 24 upregulated genes (logFC $\geq$ 2) and 85 downregulated genes (logFC $\leq$ −2), as shown in the Venn diagram (Table 1 and Fig. 1).

GO and KEGG Enrichment Analysis of the DEGs

To extract biological information from the 109 DEGs, we used DAVID for analysis (GO enrichment and KEGG analyses). The results of the GO analysis are presented from three aspects (Table 2): the biological process (BP) terms of the upregulated genes were mitotic nuclear division, regulation of attachment of spindle microtubules to kinetochore and G2/M transition of mitotic cell cycle, while the BP terms of the downregulated genes were epoxygenase P450 pathway, oxidation-reduction process, exogenous drug catabolic

process, xenobiotic metabolic process, monocarboxylic acid metabolic process, and drug metabolic process; the cellular component (CC) terms of the upregulated genes were midbody, cytoplasm, mitotic spindle, cytosol, and nucleus, while the CC terms of downregulated genes were organelle membrane, extracellular region, blood microparticle, and extracellular space; the molecular function (MF) terms of the upregulated genes were heme binding, iron ion binding, oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen, arachidonic acid epoxygenase activity, oxygen binding and monooxygenase activity, while the MF terms of downregulated genes were histone kinase activity, protein binding, protein kinase activity, and protein serine/threonine kinase activity. The results of the KEGG analysis are shown in Table 3. The upregulated genes were significantly enriched in the p53 signaling pathway and the cell cycle. The downregulated genes were significantly enriched in

Table 2. GO Enrichment Analysis of the Up- and Downregulated DEGs.

Expression	Category	Term	Count	%	p-value	FDR
Upregulated	GOTERM_BP_DIRECT	GO:0007067 ~ mitotic nuclear division	6	25	1.13E-05	0.0147121
	GOTERM_BP_DIRECT	GO:0051988 ~ regulation of attachment of spindle microtubules to kinetochore	3	12.5	3.12E-05	0.040502664
	GOTERM_BP_DIRECT	GO:0000086 ~ G2/M transition of mitotic cell cycle	4	16.67	6.34E-04	0.821644638
	GOTERM_CC_DIRECT	GO:0030496 ~ midbody	6	25	3.05E-07	3.20E-04
	GOTERM_CC_DIRECT	GO:0005737 ~ cytoplasm	17	70.83	8.67E-06	0.009070333
	GOTERM_CC_DIRECT	GO:0072686 ~ mitotic spindle	3	12.5	0.001009	1.05139872
	GOTERM_CC_DIRECT	GO:0005829 ~ cytosol	11	45.83	0.001941	2.012816159
	GOTERM_CC_DIRECT	GO:0005634 ~ nucleus	14	58.33	0.002201	2.280305423
	GOTERM_MF_DIRECT	GO:0035173 ~ histone kinase activity	2	8.33	0.004967	5.022871579
	GOTERM_MF_DIRECT	GO:0005515 ~ protein binding	18	75	0.006112	6.147950372
	GOTERM_MF_DIRECT	GO:0004672 ~ protein kinase activity	4	16.67	0.009546	9.450412521
	GOTERM_MF_DIRECT	GO:0004674 ~ protein serine/threonine kinase activity	4	16.67	0.010824	10.6521653
	GOTERM_MF_DIRECT	GO:0005524 ~ ATP binding	6	25	0.033186	29.48034821
Downregulated	GOTERM_BP_DIRECT	GO:0019373 ~ epoxygenase P450 pathway	5	6.17	1.01E-06	0.001505249
	GOTERM_BP_DIRECT	GO:0055114 ~ oxidation-reduction process	14	17.28	1.65E-06	0.0024543
	GOTERM_BP_DIRECT	GO:0042738 ~ exogenous drug catabolic process	4	4.94	1.76E-05	0.026056805
	GOTERM_BP_DIRECT	GO:0006805 ~ xenobiotic metabolic process	6	7.41	2.38E-05	0.035330156
	GOTERM_BP_DIRECT	GO:0032787 ~ monocarboxylic acid metabolic process	3	3.70	5.73E-05	0.084993128
	GOTERM_BP_DIRECT	GO:0017144 ~ drug metabolic process	4	4.94	2.23E-04	0.329889509
	GOTERM_CC_DIRECT	GO:0031090 ~ organelle membrane	8	9.88	6.75E-08	7.45E-05
	GOTERM_CC_DIRECT	GO:0005576 ~ extracellular region	23	28.40	2.98E-07	3.29E-04
	GOTERM_CC_DIRECT	GO:0072562 ~ blood microparticle	7	8.64	3.82E-05	0.042171314
	GOTERM_CC_DIRECT	GO:0005615 ~ extracellular space	15	18.52	0.001068	1.172144989
	GOTERM_MF_DIRECT	GO:0020037 ~ heme binding	8	9.88	1.72E-06	0.002191179
	GOTERM_MF_DIRECT	GO:0005506 ~ iron ion binding	8	9.88	3.58E-06	0.004576749
	GOTERM_MF_DIRECT	GO:0016705 ~ oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen	6	7.41	4.03E-06	0.005148689
	GOTERM_MF_DIRECT	GO:0016712 ~ oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen	4	4.94	3.13E-05	0.039961043
	GOTERM_MF_DIRECT	GO:0008392 ~ arachidonic acid epoxygenase activity	4	4.94	3.13E-05	0.039961043
	GOTERM_MF_DIRECT	GO:0019825 ~ oxygen binding	5	6.17	4.47E-05	0.05708859
	GOTERM_MF_DIRECT	GO:0004497 ~ monooxygenase activity	5	6.17	1.03E-04	0.131099338

DEGs: differentially expressed genes; GO: Gene Ontology; BP: biological process; CC: cellular component; MF: molecular function; FDR: false discovery rate.

Table 3. KEGG Enrichment Analysis of the Up- and Downregulated DEGs.

Expression	Pathway ID	Name	Count	%	p-value	Genes
Upregulated	hsa04115	p53 signaling pathway	3	12.5	0.001367	CCNB1, CDK1, RRM2
	hsa04110	Cell cycle	3	12.5	0.004611	CCNB1, CDK1, BUB1B
Downregulated	hsa00830	Retinol metabolism	6	7.41	4.88E-05	CYP2B6, CYP2C9, ADH4, CYP26A1, CYP2A6, CYP1A2
	hsa00232	Caffeine metabolism	3	3.70	3.95E-04	NAT2, CYP2A6, CYP1A2
	hsa00982	Drug metabolism – cytochrome P450	5	6.17	8.80E-04	CYP2B6, CYP2C9, ADH4, CYP2A6, CYP1A2
	hsa00980	Metabolism of xenobiotics by cytochrome P450	5	6.17	0.001210	CYP2B6, CYP2C9, ADH4, CYP2A6, CYP1A2
	hsa05204	Chemical carcinogenesis	5	6.17	0.001617	CYP2C9, ADH4, NAT2, CYP2A6, CYP1A2

DEGs: differentially expressed genes; KEGG: Kyoto Encyclopedia of Genes and Genomes.

retinol metabolism, caffeine metabolism, drug metabolism – cytochrome P450, metabolism of xenobiotics by cytochrome P450, and chemical carcinogenesis.

PPI Network Analysis of the DEGs and the Identification of Hub Genes

To explore the functions of these genes, we looked for proteins that interact with the DEGs in STRING and constructed a PPI network that included 69 nodes and 209 edges (Fig. 2A). Then, the hub genes of the DEGs were identified with the MCODE app from Cytoscape (Fig. 2B). In this network, we obtained 15 nodes and 105 edges; these nodes represent 15 genes (all upregulated hub genes: denticleless E3 ubiquitin protein ligase homolog (*DTL*), cyclin-dependent kinase 1 (*CDK1*), cyclin B1 (*CCNB1*), Rac GTPase-activating protein 1 (*RACGAP1*), epithelial cell transforming 2 (*ECT2*), NIMA-related kinase 2 (*NEK2*), BUB1 mitotic checkpoint serine/threonine kinase B (*BUB1B*), PDZ-binding kinase (*PBK*), DNA topoisomerase II alpha (*TOP2A*), abnormal spindle microtubule assembly (*ASPM*), hyaluronan-mediated motility receptor (*HMMR*), ribonucleotide reductase regulatory subunit M2 (*RRM2*), cyclin-dependent kinase inhibitor 3 (*CDKN3*), protein regulator of cytokinesis 1 (*PRC1*), and anillin actin-binding protein (*ANLN*) (Table 4).

Kaplan–Meier Plotter and GEPIA of the Hub Genes

After screening the 15 hub genes, we performed an OS analysis of the target genes in Kaplan–Meier plotter. The results showed that mutations in these hub genes could cause poor OS in HCC patients (Fig. 3). Subsequently, using GEPIA, we also found that the expression of the hub genes was significantly increased in tumor tissues (Fig. 4).

KEGG Analysis of the Hub Genes

After the verification of 15 hub genes using Kaplan–Meier plotter and GEPIA, KEGG pathways of these hub genes were re-analyzed via DAVID to better understand their functions; we identified four pathways associated with the 15 hub genes (Table 5). *CCNB1*, *CDK1*, and *RRM2* were enriched in the p53 signaling pathway, *CCNB1*, *CDK1*, and *BUB1B* were enriched in the cell cycle (Fig. 5), and *CCNB1* and *CDK1* were enriched in Progesterone-mediated oocyte maturation and in Oocyte meiosis ($p > 0.05$), respectively. Studied have proved that p53 signaling pathway and cell cycle play important roles in the progression of HCC[19–24]. Therefore, *CCNB1*, *CDK1*, *RRM2* and *BUB1B* may play important roles in the occurrence and development of HCC.

Discussion

HCC-related genes have yet to be identified to elucidate the underlying molecular mechanisms of cancer susceptibility, progression and prognosis. That is, a novel therapy targeting a direct regulatory mechanism that is correlated with a poor prognosis in HCC patients still needs to be identified. In this study, we selected three public microarray datasets from Asian countries, GSE121248, GSE45267, and GSE84402, and 109 DEGs were screened by GEO2 R according to the inclusion criteria between 132 HCC samples and 90 noncancerous samples. A PPI network of the DEGs was used to identify hub genes, and then the roles of these hub genes in HCC were validated via an OS analysis of patients in Kaplan–Meier plotter. The expression of these hub genes was compared between HCC and noncancerous tissues from the TCGA and GTEx. Finally, 15 hub genes related to a poor prognosis in HCC were identified: *CCNB1*, *CDK1*, *RRM2*, *BUB1B*, *DTL*, *RACGAP1*, *ECT2*, *NEK2*, *PBK*, *TOP2A*, *ASPM*, *HMMR*, *CDKN3*, *PRC1*, and *ANLN.*

In our study, among the 15 hub genes that were significantly enriched in the p53 signaling and cell cycle pathways were four potential therapeutic targets: *CCNB1*, *CDK1*, *RRM2*, and *BUB1B*. Studied have proved that cell cycle and p53 signaling pathways play important roles in development of HCC19-24. The mammalian cell cycle is controlled by regulators of the G1 to S transition such as p53, retinoblastoma, and cyclin D1 proteins. Many reports have shown that disruption of these cell cycle-related genes results in the progression of HCC19-21. And the p53 pathway is composed of a set of genes and their proteins that respond to a wide variety of stress signals. These responses to stress include cell cycle arrest, cellular senescence, or apoptosis. Moreover, the p53-regulated genes produce proteins that communicate these stress signals to adjacent cells, prevent and repair damaged DNA, and create feedback loops that regulate p53 activity and communicate with other signal transduction pathways, such as the Wnt/β-catenin, RB/INK4a, and p38 MAP pathways. Thus, the disruption of the p53 pathway has been reported in almost every type of cancer including HCC[22-24]. Based on these studies, *CCNB1*, *CDK1*, *RRM2*, and *BUB1B* may play important roles in the occurrence and development of HCC. Previous studies have reported that the CCNB1–Cdk1 complex is a key regulator of mitotic entry[25]. Importantly, Chai et al. noted that CCNB1 is highly expressed in HCC and is closely related to the poor prognosis of HCC patients, consistent with our results[26]. Gu et al. showed that CCNB1 is directly suppressed by miR-144 as a therapy targeting HCC[27]. Thus, the expression of CCNB1 is often used to estimate prognosis after treatment with anticancer drugs. CDK1 is a serine/threonine kinase and plays an important role in cell cycle progression. The inhibition of CDK1 suppresses cellular proliferation[28,29]. Its expression is also significantly higher in HCC tissue and cells; however, the real mechanism underlying the correlation between CDK1 and HCC remains unclear. RRM2 catalyzes the production of deoxynucleotide ribonucleotides in DNA synthesis and is a potential prognostic biomarker in glioma[30], breast cancer[31], and prostate cancer[32]. BUB1B, as a key mitotic spindle checkpoint, plays an important role in the development of many tumors. For example, the
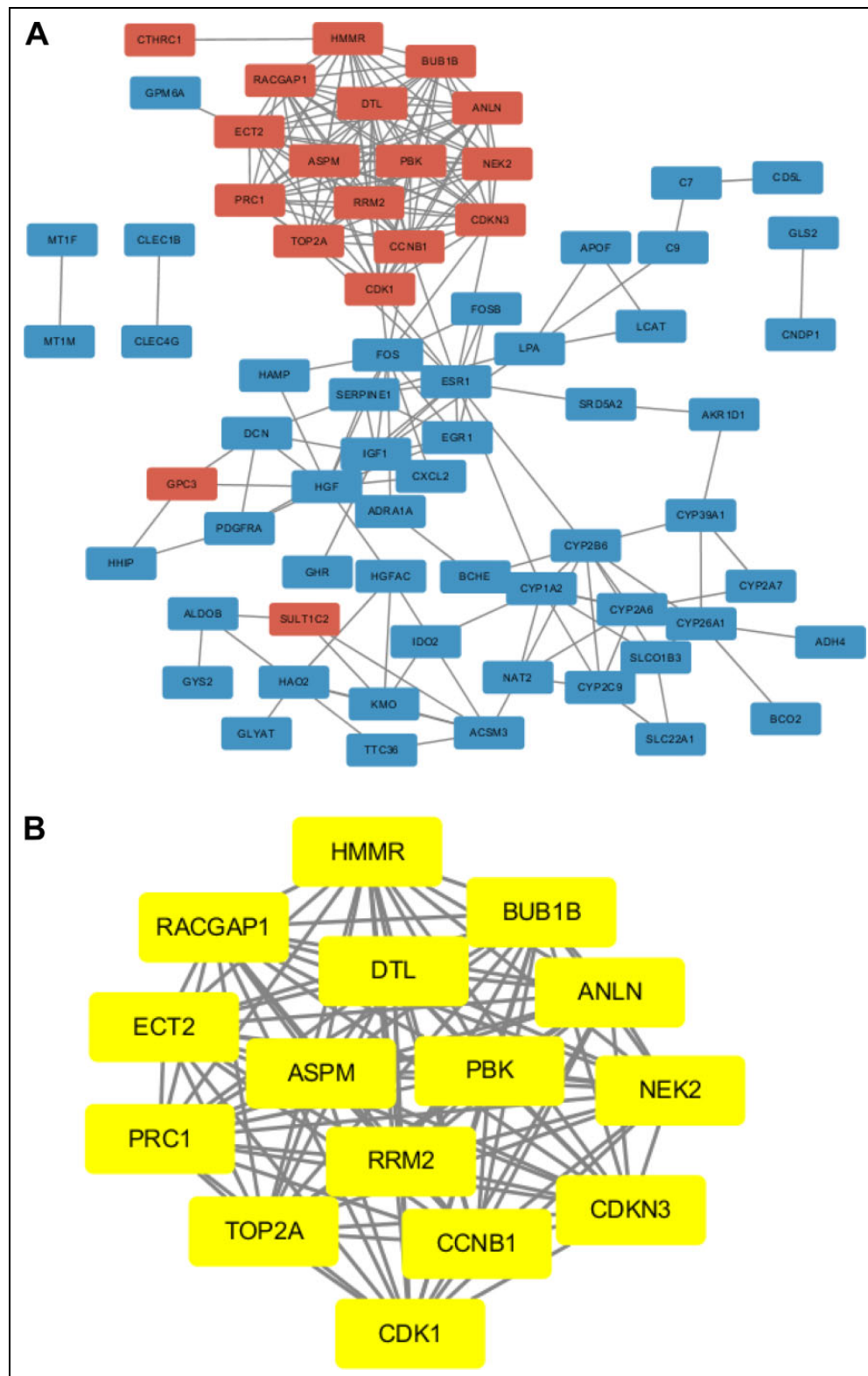
Figure 2. Identification of hub genes from the DEGs by STRING and MCODE. A. A PPI network was constructed; red nodes represent upregulated genes, and blue nodes represent downregulated genes. B. The hub genes (yellow nodes) with a degree cut-off = 2, haircut on, node score cut-off = 0.2, k-core = 2, and max. depth = 100 were screened with MCODE.

Table 4. Identification of the Hub Genes in the Protein–protein Interaction Network.

Expression	Genes
Upregulated	DTL, CDK1, CCNB1, RACGAP1, ECT2, NEK2, BUB1B, PBK, TOP2A, ASPM, HMMR, RRM2, CDKN3, PRC1, ANLN
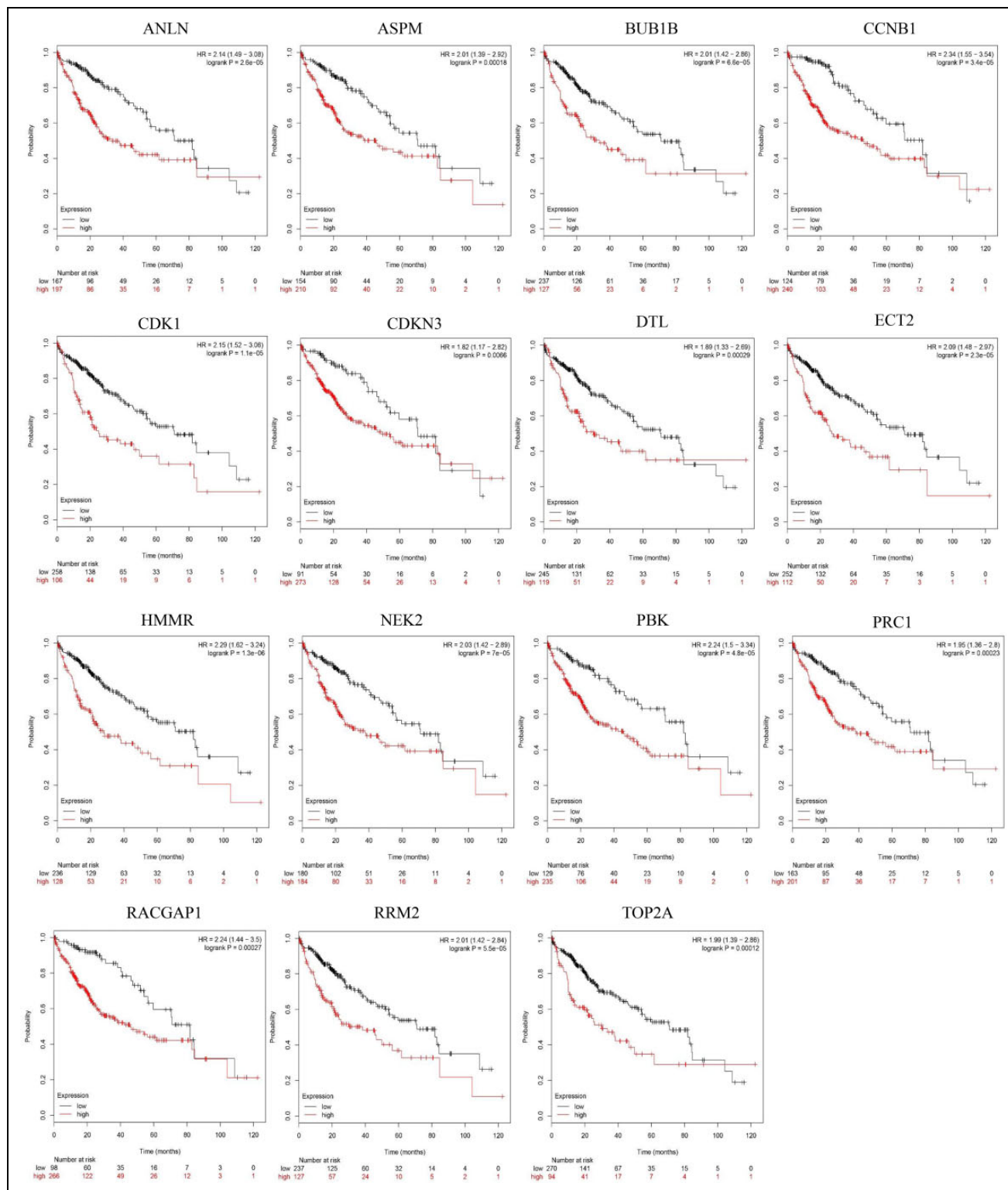
Figure 3. Overall survival analysis of the hub genes to validate survival biomarkers by Kaplan–Meier plotter. *p* < 0.05 was considered statistically significant.
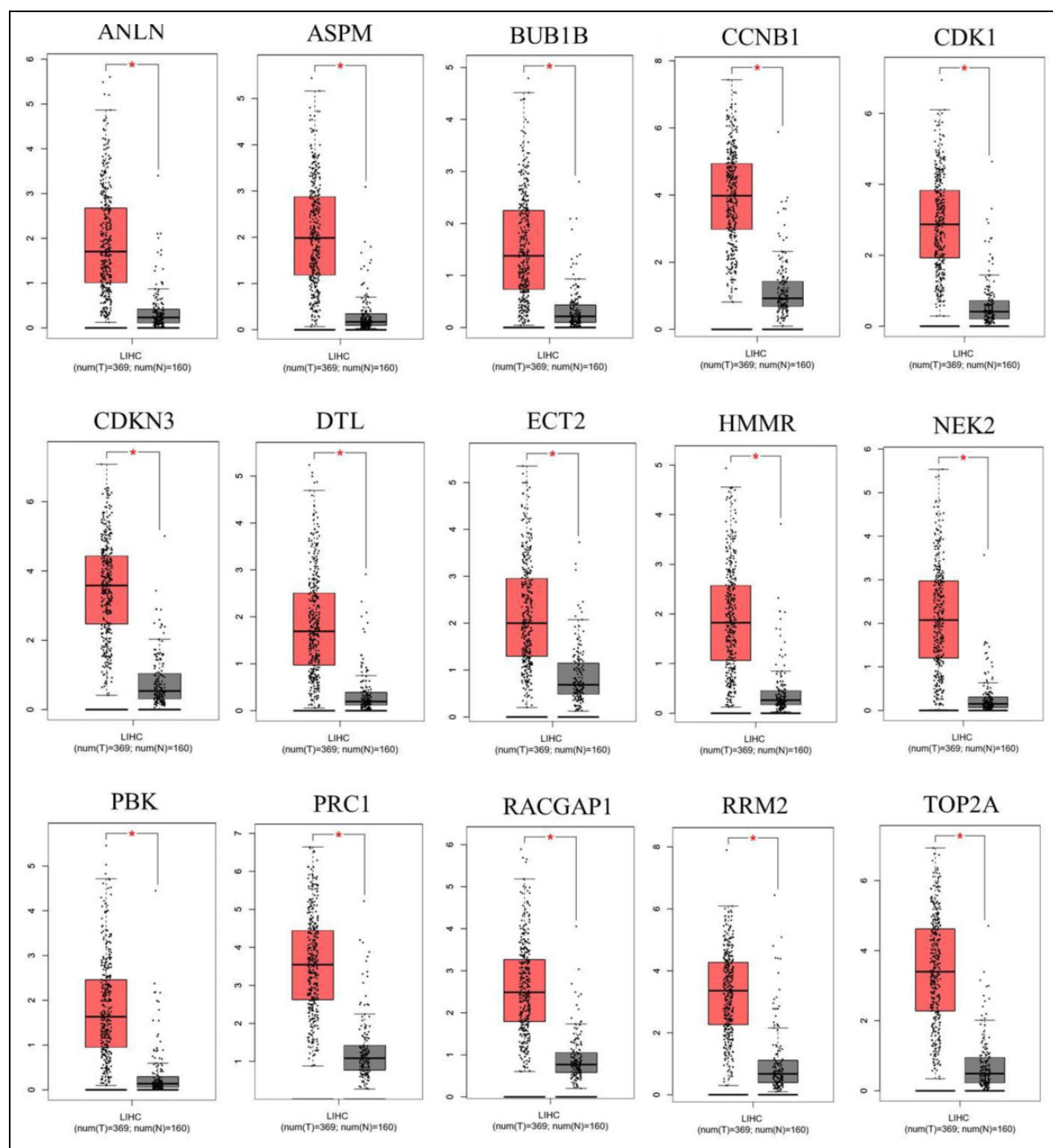
Figure 4. Expression of the hub genes was validated in 369 HCC tissues and 160 normal tissues with GEPIA. |log2FC| > 1 and *p*-value < 0.01 were considered statistically significant. Tumor tissue is shown in red, and normal tissue is shown in gray.

expression of BUB1B is increased in adrenocortical carcinomas[33], and BUB1B promotes tumor proliferation and induces radioresistance in glioblastoma[34].

We also identified other hub genes, namely, *ANLN*, *ASPM*, *CDKN3*, *DTL*, *ECT2*, *HMMR*, *NEK2*, *PBK*, *PRC1*, *RACGAP1*, and *TOP2A*, all of which encode proteins with high degrees in the PPI network. Most of these genes are

closely related to HCC. For instance, ANLN is an actin-binding protein that is essential for assembly of the cleavage furrow during cytokinesis. Lian et al.[35] and Zhang et al.[36] have shown that ANLN promotes tumor growth by decreasing apoptosis and DNA damage, and the inhibition of ANLN in liver cells blocks cytokinesis and inhibits the development of liver tumors. ASPM overexpression is a molecular marker

Table 5. KEGG Enrichment Analysis of the Hub Genes.

Pathway ID	Name	Count	%	*p*-value	Genes
hsa04115	p53 signaling pathway	3	20	5.54E-04	CCNB1, CDK1, RRM2
hsa04110	Cell cycle	3	20	0.001888704	CCNB1, CDK1, BUB1B
hsa04914	Progesterone-mediated oocyte maturation	2	13.33	0.049647728	CCNB1, CDK1
hsa04114	Oocyte meiosis	2	13.33	0.063012175	CCNB1, CDK1

KEGG: Kyoto Encyclopedia of Genes and Genomes
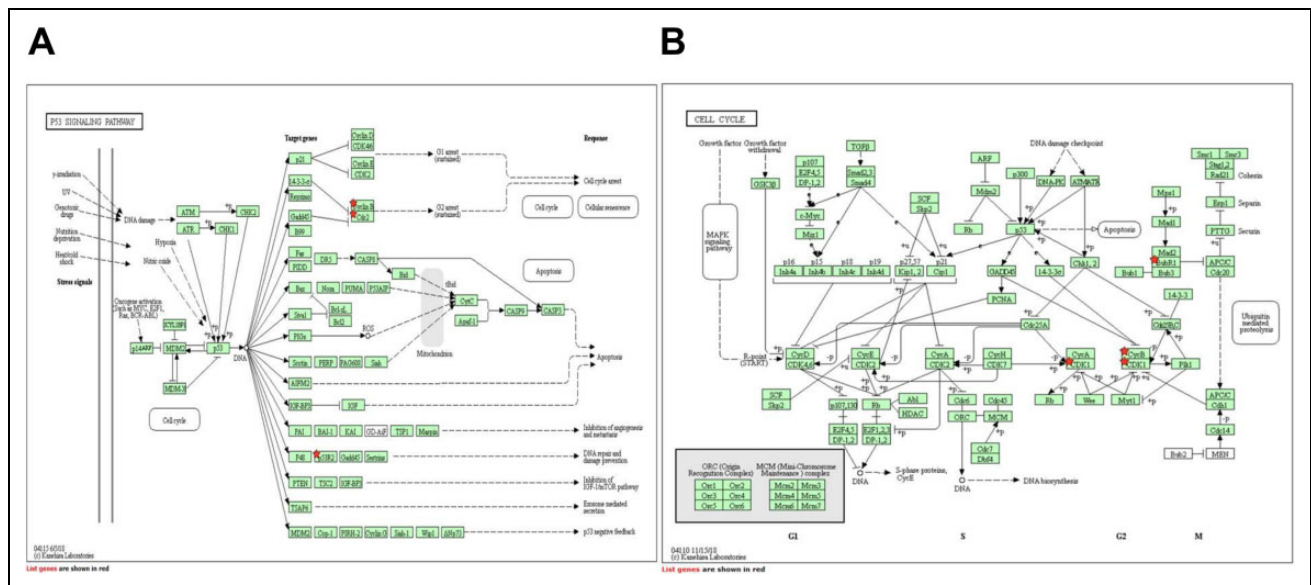
Figure 5. KEGG enrichment analysis of the hub genes. A. CCNB1, CDK1, and RRM2 are enriched in the p53 signaling pathway (Cyclin B represents CCNB1, Cdc2 represents CDK1, and p53R2 represents RRM2). B. CCNB1, CDK1, and BUB1B enriched in the cell cycle (BubR1 represents BUB1B, and CycB represents CCNB1).

that predicts the enhanced invasive/metastatic potential of HCC[37]. CDKN3 is involved in the cell cycle, and Xing et al. found that CDKN3 is frequently overexpressed in both HCC cell lines and samples and that the overexpression of CDKN3 is correlated with poor tumor differentiation and advanced tumor stage by promoting cell proliferation[38]. Chen et al. showed that PRC1 promotes early HCC recurrence and poor patient outcomes in association with the Wnt/β-catenin signaling pathway[39]. Yang et al. demonstrated that PBK promotes the metastasis of hepatocellular carcinoma by activating the ETV4-uPAR signaling pathway[40]. In addition, TOP2A contributes to the early detection and targeted therapy of a variety of cancers[41–43], including HCC[44]. However, the relationship between some of these genes, such as *NEK2*, and HCC remains unclear. NEK2 participates in the progression of multiple types of cancer, such as glioblastoma[45], adrenal cortical carcinoma[46] and myeloma[47]. Interestingly, Li et al. reported that NEK2 was overexpressed in HCC tissue and cells[48]. HCC patients with high NEK2 expression had an unfavorable prognosis, in accord with our finding. However, Fu et al. proposed that low NEK2 expression was related to a poor prognosis in HCC[49]. Therefore, the effect of

NEK2 on HCC is controversial, and its elucidation requires further evidence.

In brief, our study found that high expression of 15 hub genes was closely related to poor survival in HCC patients, indicating that their antagonism may improve the prognosis of HCC. However, the detailed mechanisms underlying the possible effects of these genes still need to be characterized *in vivo* and *in vitro*. Moreover, whether there are differences in the expression of these genes in different stages of HCC or different diseases also needs to be further studied and discussed.

Ethical Approval

Ethical Approval is not applicable for this article.

Statement of Human and Animal Rights

This article does not contain any studies with human or animal subjects.

Statement of Informed Consent

There are no human subjects in this article and informed consent is not applicable.

Declaration of Conflicting Interests

Funding

ORCID iD

Zengyuan Zhou https://orcid.org/0000-0001-7216-3165

Supplemental Material

Supplemental material for this article is available online.

References

1. Islami F, Goding Sauer A, Miller KD, Siegel RL, Fedewa SA, Jacobs EJ, McCullough ML, Patel AV, Ma J, Soerjomataram I, Flanders WD, et al. Proportion and number of cancer cases and deaths attributable to potentially modifiable risk factors in the United States. CA Cancer J Clin. 2018;68(1):31–54.

2. Degasperi E, Colombo M. Distinctive features of hepatocellular carcinoma in non-alcoholic fatty liver disease. Lancet Gastroenterol Hepatol. 2016;1(2):156–164.

3. Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. Lancet. 2012;379(9822):1245–1255.

4. Estes C, Anstee QM, Arias-Loste MT, Bantel H, Bellentani S, Caballeria J, Colombo M, Craxi A, Crespo J, Day CP, Eguchi Y, et al. Modeling NAFLD disease burden in China, France, Germany, Italy, Japan, Spain, United Kingdom, and United States for the period 2016-2030. J Hepatol. 2018;69(4): 896–904.

5. Kulik L, El-Serag HB. Epidemiology and management of hepatocellular carcinoma. Gastroenterology. 2019;156(2): 477–491.e1.

6. Nault JC, De Reyniès A, Villanueva A, Calderaro J, Rebouissou S, Couchy G, Decaens T, Franco D, Imbeaud S, Rousseau F, Azoulay D, et al. A hepatocellular carcinoma 5-gene score associated with survival of patients after liver resection. Gastroenterology. 2013;145(1):176–187.

7. Gores GJ.Decade in review-hepatocellular carcinoma: HCC-subtypes, stratification and sorafenib. Nat Rev Gastroenterol Hepatol. 2014;11(11):645–647.

8. Khemlina G, Ikeda S, Kurzrock R. The biology of hepatocellular carcinoma: implications for genomic and immune therapies. Mol Cancer. 2017;16(1):149.

9. Díaz-Mochón JJ, Tourniaire G, Bradley M. Microarray platforms for enzymatic and cell-based assays. Chem Soc Rev. 2007;36(3):449–457.

10. Wang SM, Ooi LL, 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.

11. Chen CL, Tsai YS, Huang YH, Liang YJ, Sun YY, Su CW, Chau GY, Yeh YC, Chang YS, Hu JT, Wu JC. Lymphoid enhancer factor 1 contributes to hepatocellular carcinoma progression through transcriptional regulation of epithelial-mesenchymal transition regulators and stemness genes. Hepatol Commun. 2018;2(11):1392–1407.

12. Wang H, Huo X, Yang XR, He J, Cheng L, Wang N, Deng X, Jin H, Wang N, Wang C, Zhao F, et al. STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4. Mol Cancer. 2017;16(1):136.

13. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, Marshall KA, Phillippy KH, Sherman PM, Holko M, Yefanov A, et al. NCBI GEO: archive for functional genomics data sets–update. Nucleic Acids Res. 2013;41(Database issue):D991–D995.

14. Mazandu GK, Chimusa ER, Mulder NJ. Gene ontology semantic similarity tools: survey on features and challenges for biological knowledge discovery. Brief Bioinform. 2017;18(5): 886–901.

15. Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. Nucleic Acids Res. 2017;45(D1):D353–D361.

16. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, Lin J, Minguez P, Bork P, von Mering C, Jensen LJ. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2013; 41(Database issue):D808–D815.

17. Bader GD, Hogue CWV. An automated method for finding molecular complexes in large protein interaction networks. BMC Bioinformatics. 2003;4:2.

18. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res. 2017;45(W1): W98–W102.

19. Farazi PA, DePinho RA. Hepatocellular carcinoma pathogenesis: from genes to environment. Nat Rev Cancer. 2006;6(9): 674–687.

20. Teufel A, Staib F, Kanzler S, Weinmann A, Schulze-Bergkamen H, Galle PR. Genetics of hepatocellular carcinoma. World J Gastroenterol. 2007;13(16):2271–2282.

21. Joo M, Kang YK, Kim MR, Lee HK, Jang JJ. Cyclin D1 over-expression in hepatocellular carcinoma. Liver. 2001;21(2): 89–95.

22. Levine AJ, Hu W, Feng Z. The P53 pathway: what questions remain to be explored? Cell Death Differ. 2006;13(6): 1027–1036.

23. Nishida N, Goel A. Genetic and epigenetic signatures in human hepatocellular carcinoma: a systematic review. Curr Genomics. 2011;12(2):130–137.

24. Meng X, Franklin DA, Dong J, Zhang Y. MDM2-p53 pathway in hepatocellular carcinoma. Cancer Res. 2014;74(24): 7161–7167.

25. Nakayama Y, Yamaguchi N. Role of cyclin B1 levels in DNA damage and DNA damage-induced senescence. Int Rev Cell Mol Biol. 2013;305:303–337.

26. Chai N, Xie HH, Yin JP, Sa KD, Guo Y, Wang M, Liu J, Zhang XF, Zhang X, Yin H, Nie YZ, et al. FOXM1 promotes proliferation in human hepatocellular carcinoma cells by transcriptional activation of CCNB1. Biochem Biophys Res Commun. 2018;500(4):924–929.

27. Gu J, Liu X, Li J, He Y. MicroRNA-144 inhibits cell proliferation, migration and invasion in human hepatocellular carcinoma by targeting CCNB1. Cancer Cell Int. 2019;19:15.

28. Lv Y, Xiao J, Liu J, Xing F. E2F8 is a potential therapeutic target for hepatocellular carcinoma. J Cancer. 2017;8(7): 1205–1213.

29. Zhang Y, Huang W, Ran Y, Xiong Y, Zhong Z, Fan X, Wang Z, Ye Q. miR-582-5p inhibits proliferation of hepatocellular carcinoma by targeting CDK1 and AKT3. Tumour Biol. 2015; 36(11):8309–8316.

30. Sun H, Yang B, Zhang H, Song J, Zhang Y, Xing J, Yang Z, Wei C, Xu T, Yu Z, Xu Z, et al. RRM2 is a potential prognostic biomarker with functional significance in glioma. Int. J Biol Sci. 2019;15(3):533–543.

31. Chen WX, Yang LG, Xu LY, Cheng L, Qian Q, Sun L, Zhu YL. Bioinformatics analysis revealing prognostic significance of RRM2 gene in breast cancer. Biosci Rep. 2019;39(4): BSR20182062.

32. Mazzu YZ, Armenia J, Chakraborty G, Yoshikawa Y, Coggins SA, Nandakumar S, Gerke TA, Pomerantz MM, Qiu X, Zhao H, Atiq M, et al. A novel mechanism driving poor-prognosis prostate cancer: overexpression of the DNA repair gene, ribonucleotide reductase small subunit M2 (RRM2). Clin Cancer Res. 2019;25(14):4480–4492.

33. Passaia BDS, Dias MH, Kremer JL, Antonini SRR, de Almeida MQ, Fragoso MCBV, Lotfi CFP. TCF21/POD-1, a transcritional regulator of SF-1/NR5A1, as a potential prognosis marker in adult and pediatric adrenocortical tumors. Front Endocrinol (Lausanne). 2018;9:38.

34. Ma Q, Liu Y, Shang L, Yu J, Qu Q. The FOXM1/BUB1B signaling pathway is essential for the tumorigenicity and radioresistance of glioblastoma. Oncol Rep. 2017;38(6): 3367–3375.

35. Lian YF, Huang YL, Wang JL, Deng MH, Xia TL, Zeng MS, Chen MS, Wang HB, Huang YH. Anillin is required for tumor growth and regulated by miR-15a/miR-16 -1 in HBV-related hepatocellular carcinoma. Aging (Albany NY). 2018;10(8): 1884–1901.

36. Zhang S, Nguyen LH, Zhou K, Tu HC, Sehgal A, Nassour I, Li L, Gopal P, Goodman J, Singal AG, Yopp A, et al. Knockdown of anillin actin binding protein blocks cytokinesis in hepatocytes and reduces liver tumor development in mice without affecting regeneration. Gastroenterology. 2018;154(5): 1421–1434.

37. Lin SY, Pan HW, Liu SH, Jeng YM, Hu FC, Peng SY, Lai PL, Hsu HC. ASPM is a novel marker for vascular invasion, early recurrence, and poor prognosis of hepatocellular carcinoma. Clin Cancer Res. 2008;14(15):4814–4820.

38. Xing C, Xie H, Zhou L, Zhou W, Zhang W, Ding S, Wei B, Yu X, Su R, Zheng S. Cyclin-dependent kinase inhibitor 3 is overexpressed in hepatocellular carcinoma and promotes tumor cell proliferation. Biochem Biophys Res Commun. 2012;420(1): 29–35.

39. Chen J, Rajasekaran M, Xia H, Zhang X, Kong SN, Sekar K, Seshachalam VP, Deivasigamani A, Goh BK, Ooi LL, Hong W, et al. The microtubule-associated protein PRC1 promotes early recurrence of hepatocellular carcinoma in association with the Wnt/β-catenin signalling pathway. Gut. 2016;65(9): 1522–1534.

40. Yang QX, Zhong S, He L, Jia XJ, Tang H, Cheng ST, Ren JH, Yu HB, Zhou L, Zhou HZ, Ren F, et al. PBK overexpression promotes metastasis of hepatocellular carcinoma via activating ETV4-uPAR signaling pathway. Cancer Lett. 2019;452: 90–102.

41. Labbé DP, Sweeney CJ, Brown M, Galbo P, Rosario S, Wadosky KM, Ku SY, Sjöström M, Alshalalfa M, Erho N, Davicioni E, et al. TOP2A and EZH2 provide early detection of an aggressive prostate cancer subgroup. Clin Cancer Res. 2017;23(22):7072–7083.

42. Chen JR, Chien HP, Chen KS, Hwang CC, Chen HY, Yeh KY, Hsieh TY, Chang LC, Hsu YC, Lu RJ, Hua CC. Amplification of HER2 and TOP2A and deletion of TOP2A genes in a series of Taiwanese breast cancer. Medicine (Baltimore). 2017;96(2): e5582.

43. Aumayr K, Klatte T, Neudert B, Birner P, Shariat S, Schmidinger M, Susani M, Haitel A. HER2 and TOP2A gene amplification and protein expression in upper tract urothelial carcinomas. Pathol Oncol Res. 2018;24(3):575–581.

44. Panvichian R, Tantiwetrueangdet A, Angkathunyakul N, Leelaudomlipi S. TOP2A amplification and overexpression in hepatocellular carcinoma tissues. Biomed Res Int. 2015; 2015:381602.

45. Wang J, Cheng P, Pavlyukov MS, Yu H, Zhang Z, Kim SH, Minata M, Mohyeldin A, Xie W, Chen D, Goidts V, et al. Targeting NEK2 attenuates glioblastoma growth and radioresistance by destabilizing histone methyltransferase EZH2. J Clin Invest. 2017;127(8):3075–3089.

46. Mete O, Gucer H, Kefeli M, Asa SL. Diagnostic and prognostic biomarkers of adrenal cortical carcinoma. Am J Surg Pathol. 2018;42(2):201–213.

47. Franqui-Machin R, Hao M, Bai H, Gu Z, Zhan X, Habelhah H, Jethava Y, Qiu L, Frech I, Tricot G, Zhan F. Destabilizing NEK2 overcomes resistance to proteasome inhibition in multiple myeloma. J Clin Invest. 2018;128(7):2877–2893.

48. Li G, Zhong Y, Shen Q, Zhou Y, Deng X, Li C, Chen J, Zhou Y, He M. NEK2 serves as a prognostic biomarker for hepatocellular carcinoma. Int J Oncol. 2017;50(2):405–413.

49. Fu L, Liu S, Wang H, Ma Y, Li L, He X, Mou X, Tong X, Hu Z, Ru G. Low expression of NEK2 is associated with hepatocellular carcinoma progression and poor prognosis. Cancer Biomark. 2017;20(1):101–106.