

# Identification of Hub Genes Associated with Hepatocellular Carcinoma Prognosis by Bioinformatics Analysis

Xi Zhang\*, Xiaojun Luo\*, Wenbin Liu, Ai Shen<sup>#</sup>

Department of Hepatobiliary and Pancreatic Tumor Center, Chongqing University Cancer Hospital, Chongqing, China  
Email: 15223434456@sina.cn, luocom163@163.com, chevymoss@163.com, #shenai200808@163.com

**How to cite this paper:** Zhang, X., Luo, X.J., Liu, W.B. and Shen, A. (2021) Identification of Hub Genes Associated with Hepatocellular Carcinoma Prognosis by Bioinformatics Analysis. *Journal of Cancer Therapy*, 12, 186-207.

<https://doi.org/10.4236/jct.2021.124019>

**Received:** March 23, 2021

**Accepted:** April 26, 2021

**Published:** April 29, 2021

Copyright © 2021 by author(s) and Scientific Research Publishing Inc.  
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

**Objective:** This study aimed to identify hub genes that are associated with hepatocellular carcinoma (HCC) prognosis by bioinformatics analysis. **Methods:** Data were collected from the Gene Expression Omnibus (GEO) and The Cancer Genome Atlas (TCGA) liver HCC datasets. The robust rank aggregation algorithm was used in integrating the data on differentially expressed genes (DEGs). Online databases DAVID 6.8 and REACTOME were used for gene ontology and pathway enrichment analysis. R software version 3.5.1, Cytoscape, and Kaplan-Meier plotter were used to identify hub genes. **Results:** Six GEO datasets and the TCGA liver HCC dataset were included in this analysis. A total of 151 upregulated and 245 downregulated DEGs were identified. The upregulated DEGs most significantly enriched in the functional categories of cell division, chromosomes, centromeric regions, and protein binding, whereas the downregulated DEGs most significantly enriched in the epoxygenase P450 pathway, extracellular region, and heme binding, with respect to biological process, cellular component, and molecular function analysis, respectively. Upregulated DEGs most significantly enriched the cell cycle pathway, whereas downregulated DEGs most significantly enriched the metabolism pathway. Finally, 88 upregulated and 40 downregulated genes were identified as hub genes. The top 10 upregulated hub DEGs were *CDK1*, *CCNB1*, *CCNB2*, *CDC20*, *CCNA2*, *AURKA*, *MAD2L1*, *TOP2A*, *BUB1B* and *BUB1*. The top 10 downregulated hub DEGs were *ESR1*, *IGF1*, *FTCD*, *CYP3A4*, *SPP2*, *C8A*, *CYP2E1*, *TAT*, *P9* and *CYP2C9*. **Conclusions:** This study identified several upregulated and downregulated hub genes that are associated with the prognosis of HCC patients. Verification of these results using *in vitro* and *in vivo* studies is warranted.

\*The authors contributed equally to this work.

<sup>#</sup>Corresponding author.

---

## Keywords

Hepatocellular Carcinoma, Hub Genes, Bioinformatics, Robust Rank Aggregation, Differentially Expressed Genes

---

## 1. Introduction

Hepatocellular carcinoma (HCC) is one of the six most common cancers and the third leading cause of cancer-related death [1]. The most effective treatment for HCC is curative resection, which includes liver transplantation and hepatectomy [2]. However, only 20% of the HCC patients are candidates for curative hepatectomy and HCC often recurs shortly after surgery [2]. Palliative treatment is the main treatment modality for most HCC patients [2]. However, transarterial chemoembolization is only a regional therapy for intrahepatic tumors [2], and it cannot be used for extrahepatic metastatic tumors, including lung metastasis, bone metastasis, or circulating tumor cells. Hence, local therapy combined with systemic therapy may be an ideal scheme. Although most HCC are caused by viral infections, anti-viral therapy is somehow effective [3], but it does not directly prevent the occurrence and development of liver tumors. Therefore, target therapy for tumorigenesis may be a promising systematic treatment. However, current target therapies have limited effects on the prognosis of HCC patients. Although sorafenib could improve the survival time of advanced HCC patients, it has not substantially changed the outcome of patients with HCC, and the median survival time of patients treated with sorafenib is only three months longer than with palliative treatment [4]. Additionally, the application of sorafenib post-hepatectomy does not achieve the expected benefits [5]. The median disease-free survival time was 8.5 (2.9 - 19.5) and 8.4 (2.9 - 19.8) months in the sorafenib and placebo groups, respectively [5].

Although many researchers focused on studying the mechanism of HCC tumorigenesis [6] [7] [8], the mechanism of hepatocellular tumorigenesis is still not been fully elucidated, and an effective therapeutic target is still needed to improve the prognosis of HCC. In recent decades, due to the rapid development of high-throughput sequencing and the wide application of gene chips, extensive studies elucidating the mechanism of HCC tumorigenesis have been conducted, and bioinformatics analysis has helped in identifying key genes in hepatocellular carcinogenesis [7] [9] [10]. Shen *et al.* reported that *TOP2A*, *NDC80*, *FOXM1*, *HMMR*, *KNTC1*, *PTTG1*, *FEN1*, *RFC4*, *SMC4*, and *PRC1* are the top 10 core genes associated with HCC tumorigenesis [10]. Chen *et al.* reported that *TOP2A*, *CDC20*, *MAD2L1*, *BUB1B*, *RFC4*, *CCNB1*, *CDKN3*, *CCNB2*, *TPX2*, and *FEN1* are the 10 hub genes in hepatitis B virus-related HCC [9]. However, some datasets in these studies consisted of less than 10 cases, and the Venn diagram was used to integrate their gene expression data. These two factors may miss several important genes during extraction of DEGs [9] [10]. Therefore, to identify the hub gene of HCC

and to achieve more reliable results, this study also conducted bioinformatics analysis based on several datasets and each dataset consisting of >30 cases. In addition, robust rank aggregation (RRA) algorithm was used to integrate the expression data of various datasets to identify differentially expressed genes (DEGs).

## 2. Materials and Methods

### 2.1. Microarray Datasets

Gene expression data were retrieved from Gene Expression Omnibus (GEO) (<https://www.ncbi.nlm.nih.gov/geo>) and The Cancer Genome Atlas (TCGA) liver HCC datasets (<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>). Only datasets with cases > 30 were included in the analysis. Only datasets consisting of HCC tissues and adjacent non-tumor tissues were included in the analysis.

### 2.2. Identification of the DEGs

Identification of DEGs was performed with R software ver. 3.5.1 using the “limmapackage”. The  $\log_2$ |fold change (FC)| of gene expression in tumor tissues compared to adjacent non-tumor tissues was calculated. Gene expression difference with  $|\log FC| > 1.0$  and adjusted p value < 0.05 were regarded as DEGs for each included dataset. However, DEGs among these datasets were calculated using R software based on “RobustRankAggreg” package [11]. Integrate genes with significance score lower than 0.05 was regarded as DEGs [11]. These DEGs were verified using Oncomine (<https://www.oncomine.org/resource/main.html>) and GEPIA (<http://gepia.cancer-pku.cn/>).

### 2.3. Gene Ontology and Pathway Enrichment Analysis of DEGs

Online database DAVID 6.8 (<https://david.ncifcrf.gov/>) was used for Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis of DEGs. GO analysis consists of biological process (BP), cellular component (CC), and molecular function (MF) analysis. The online database Reactome was also used for pathway enrichment analysis of DEGs (<http://www.reactome.org>).

### 2.4. Identification of Potential Hub DEGs by Protein-Protein Interaction (PPI) Network Analysis

A PPI network was constructed using online STRING database (<http://string-db.org>). Cytoscape software version 3.7.1 was used to identify potential hub genes. DEGs with interaction degrees  $\geq 10$  were regarded as potential hub genes. MCODE plug-ins were used to explore closely related functional modules in the PPI networks. The degree cutoff of module network scoring was set to 2. The node score cutoff, K-core, and maximum depth of module finding were set to 0.2, 2, and 100, respectively.

### 2.5. Identification of Hub DEGs by Survival Analysis

The clinical data were collected from the TCGA liver HCC dataset. Clinical data

from a total of 364 patients were included in the survival analysis. The survival results were calculated using Kaplan Meier plotter (<http://kmplot.com/analysis/>). The Kaplan Meier method was used in survival analysis. These results were verified using the online database GEPIA (<http://gepia.cancer-pku.cn/>).

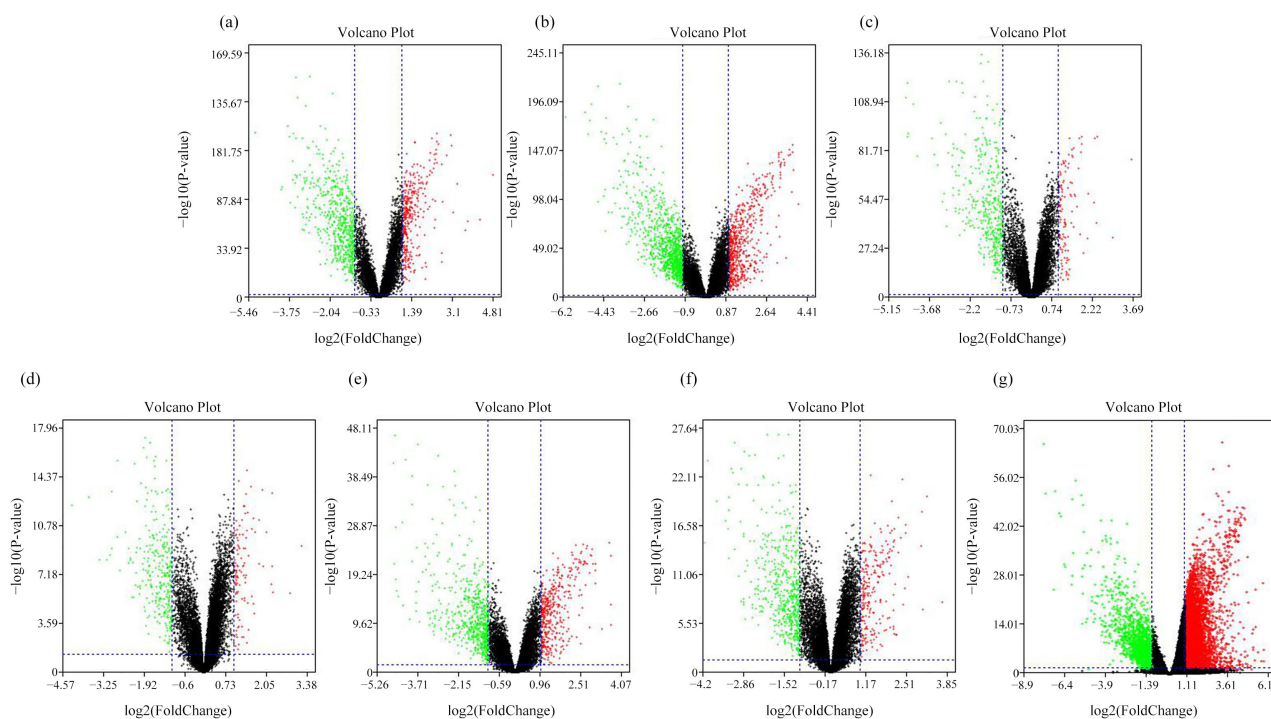
### 3. Results

#### 3.1. Data Characteristics

Six GEO datasets (GSE14520, GSE25097, GSE36376, GSE57957, GSE76427, and GSE121248) and the TCGA liver HCC dataset were analyzed in this study. A total of 1328 tumor tissues and 834 adjacent non-tumor tissues were included in the analysis. The characteristics of the included datasets were detailed in **Table 1**. All datasets included tumor tissues and adjacent non-tumor tissues.

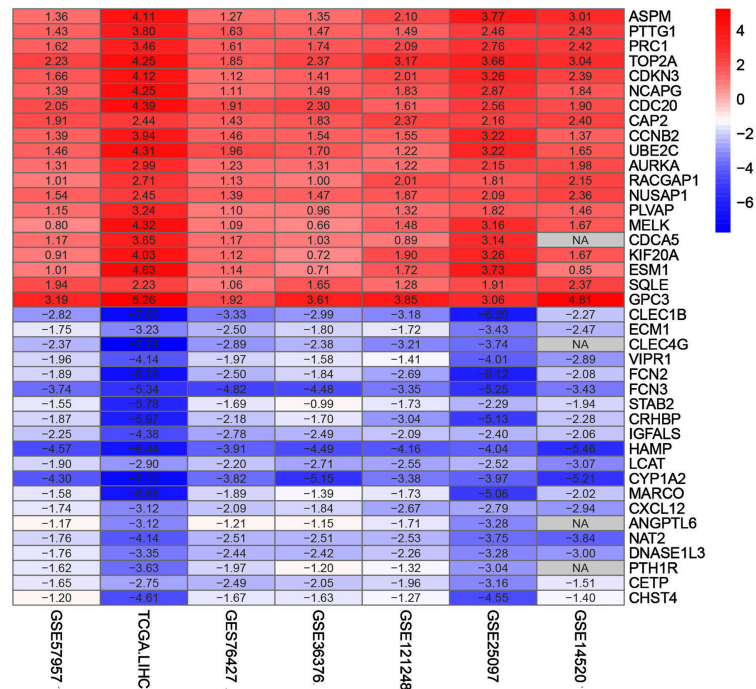
#### 3.2. Identification of DEGs

Using the R software, DEGs in each dataset were identified. The gene expression data in each dataset are shown using a volcano map (**Figure 1**). The RRA algorithm identified significant DEGs among these datasets. A total of 151 upregulated and 245 downregulated genes were identified (**Table 2**). The top 20 upregulated and downregulated DEGs are shown in **Figure 2**.



**Figure 1.** Volcano plot of differentially expressed genes in each included dataset ((a) GSE14520 dataset; (b) GSE25097 dataset; (c) GSE36376 dataset; (d) GSE57957 dataset; (e) GSE84005 dataset; (f) GSE121248 dataset; (g) TCGA\_LIHC dataset). The differentially expressed genes were identified by “limma” package of the R software for each dataset. FC, fold change. Red dots, upregulated expressed genes based on  $\log_2|\text{fold change}| > 1.0$  and adjusted P-value  $< 0.05$ . Green dots, downregulated expressed genes based on  $\log_2|\text{fold change}| > 1.0$  and adjusted P-value  $< 0.05$ . Black dots, expressed genes with no significant difference.





**Figure 2.** Top 20 upregulated and downregulated genes identified in the datasets using the RRA algorithm. Red background, upregulated genes. Green background, downregulated genes. The numbers are the fold change of the differentially expressed genes. NA, not available.

**Table 1.** Baseline characteristics of included datasets.

Datasets	Platforms	Tumor tissue (cases)	Adjacent non-tumor tissue (cases)
GSE14520	GPL3921 Affymetrix HT Human Genome U133A Array	225	220
GSE25097	GPL10687 Rosetta/Merck Human RSTA Affymetrix 1.0 microarray	268	243
GSE36376	GPL10558 Illumina HumanHT-12 V4.0 expression beadchip	240	193
GSE57957	GPL10558 Illumina HumanHT-12 V4.0 expression beadchip	39	39
GSE76427	GPL10558 Illumina HumanHT-12 V4.0 expression beadchip	115	52
GSE121248	GPL570 Affymetrix Human Genome U133 Plus 2.0 Array	70	37
TCGA-LIHC	Illumina HiSeq	371	50

GSE, Gene Expression Omnibus. TCGA-LIHC, The Cancer Genome Atlas-Liver Hepatocellular Carcinoma.

**Table 2.** Differentially expressed genes identified by robust rank aggregation algorithm via R software.

Differentially expressed genes
Upregulated (151)
ASPM PTTG1 PRC1 TOP2A CDKN3 NCAPG CDC20 CAP2 CCNB2 UBE2C AURKA RACGAP1 NUSAP1 PLVAP MELK CDCA5 KIF20A ESM1 SQLC GPC3 ILF2 UBE2T FAM189B PBK ACSL4 CENPF NEK2 MCM2 HMMR AKR1C3 KIFC1 GMNN CKAP2L ANXA2 CCNB1 MDK DTL FAM83D SSR2 CD34 RRM2 PTTG3P KIF4A TOMM40L COL15A1 SPINK1 TTK RFC4 PEA15 AKR1B10 RFX5 THY1 BUB1B FOXM1 PODXL ECT2 TP53I3 EZH2 NDC80 CCNA2 ZWINT COL4A1 CDK1 TXNRD1 MCM3 PSMD4 CCL20 ANLN TKT AURKB TK1 TRIP13 S100A10 MCM6 KIF2C NUF2 H2AFZ AKR1B15 CCDC34 FLVCR1 ITGA6 MND1 DLGAP5 SKA1 CDCA3 TCF19 NME1 CCT3 NDUFA4L2 CENPE MSH2 NQO1 S100P RAD51AP1 KIF18B CDC25C EXO1 CDC6 HJURP REG3A SPP1 GABRD E2F8 LCN2 BIRC5 SNRPB LAPTM4B SPC25 GINS1 MAD2L1 SLC26A6 ATAD2 LAMC1 SKA3 TUBA1C CKAP2 SF3B4 ACLY TPX2 HELLS MCM4 CDC7 UHRF1 LRRC1 C6orf173 COX7B2 STMN1 CKS2 HIGD1B CDKN2C CCNE2 CENPA THBS4 SGOL2 C5orf34 MGC4677 HSP90AB1 TFRC KPNA2 KIAA0101 PEG10 BUB1 ALG1L STIL MKI67 FABP5L2 CKS1B TARBP1 BRK1 EME1 SCAMP3

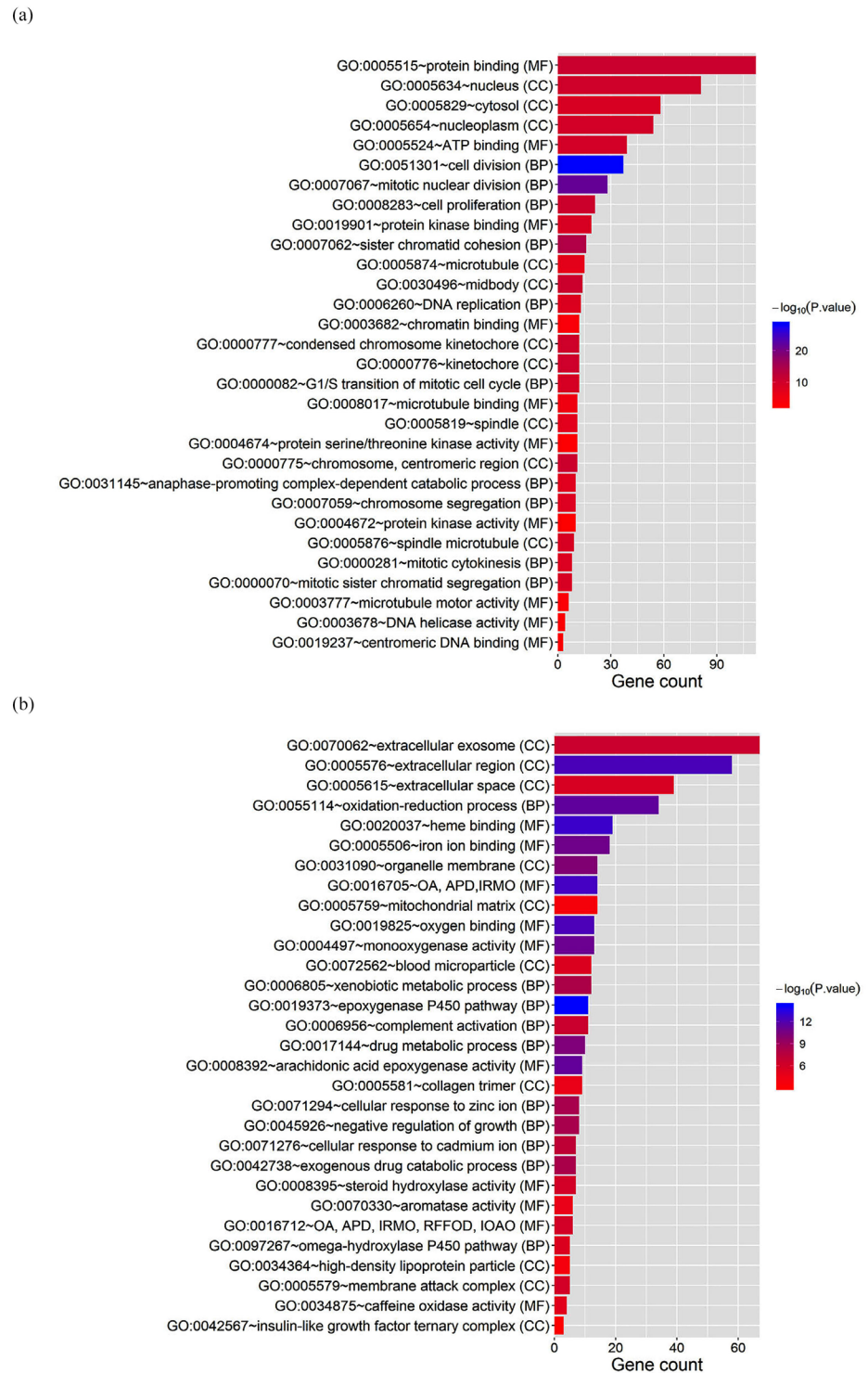
## Continued

Downregulated (245)	CLEC1B ECM1 CLEC4G VIPR1 FCN2 FCN3 STAB2 CRHBP IGFALS HAMP LCAT CYP1A2 MARCO CXCL12 ANGPTL6 NAT2 DNASE1L3 PTH1R CETP CHST4 RND3 COLEC10 CYP39A1 CXCL14 FOS KCNN2 SRPX ACADS ESR1 TTC36 CNDP1 CFP CPEB3 DBH PAMR1 GSTZ1 GPR128 OIT3 CLEC4M LY6E CYP26A1 DPT ID1 FOSB MT1F DCN APOF TMEM27 EGR1 KMO ZFP36 C8orf4 LYVE1 KBTBD11 ACSM3 AADAT ZGPAT SRD5A2 SOCS2 ASS1 GADD45B C7 SLC01B3 IGFBP3 SHBG RCAN1 MT1X CYR61 INMT MT1M HGFAC ENO3 ANXA10 C1RL ETFDH MT1G STEAP3 RSP03 GCDH FBP1 SKAP1 PZP JMJD5 RDH5 MT1H COLEC11 MYOM2 NNMT GHR MT2A RCL1 CYP4A11 PROZ KLKB1 ADAMTS13 SLC22A1 C1QTNF1 HAO2 GABARAPL1 FXYD1 MT1E CYP3A4 C9 ADAMTSL2 MS4A6A GLS2 ALDH2 PLAC8 AGL MSRA ACAA1 CCBE1 NDRG2 RDH16 FAM13A PHLDA1 BMPER STARD5 CYP2B6 ADRA1A FAM134B HHIP AKR7A3 ST3GAL6 TACSTD2 BCHE SPP2 PCDH24 CYP2C9 C8A SIGIRR CYP2C8 ACACB ABCA8 BBOX1 DNAJC12 ATOH8 HBA2 GLYAT IL1RAP ACSL1 PBLD AFM LIFR TDO2 CDC37L1 ALDH8A1 IDO2 MBL2 OGDHL GBA3 CXCL2 PANK1 GPM6A SLC27A5 FTCD EPHX2 CTH C6 PHGDH PEMT CD5L CYP4V2 MAT1A PCK1 ITLN1 CLRN3 CYP2E1 CYP2A6 C8B GNMT EPB41L4B SORL1 BHMT GCH1 ADH4 CCL14 FAHD2A NPY1R LRRN3 HSD17B13 ACAA2 SLC7A2 NR1I2 LPA F9 CCRN4L ARHGAP10 SLC10A1 C1R ASFG FOXO1 FLJ21986 THRSP GADD45G IGF1 ADH1B GLYATL1 TRIB1 IGJ TBXA2R TAT SDS CYP4F2 HBB GPD1 HSD11B1 MFAP4 HSD17B2 PEX11G AXUD1 ATF5 CSRN1 WDR72 SAA4 ALDH6A1 CIDEA ASPA CBR4 C14orf68 VNN1 IRF8 MT1A PLA2G16 MRO TMEM45A LIPC PDGFRA JCLN CCL23 SLC19A3 PPAP2B BCO2 OLFML3 CYP2A7 PGLYRP2 HPX COL6A6 CYP4F12 RANBP3L NPC1L1 HMGCL ALPL AGXT2 GPR182
------------------------	-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

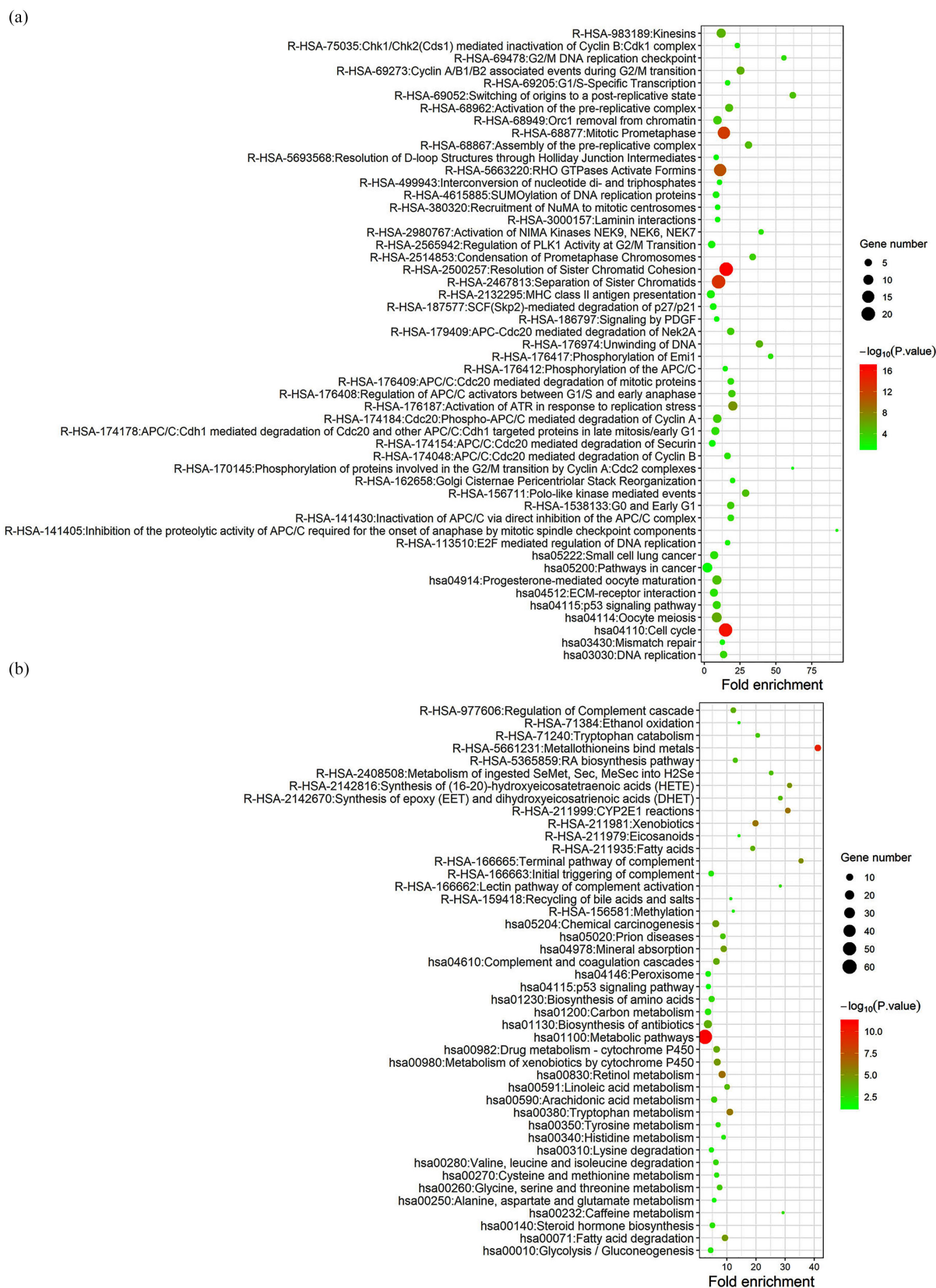
### 3.3. GO and Pathway Enrichment Analysis of DEGs

After GO enrichment analysis for upregulated DEGs, a total of 74, 37, and 26 items with *P* values <0.05 were identified in BP, CC, and MF enrichment analysis, respectively, while for downregulated DEGs, a total of 87, 21, and 34 items with *P* values <0.05 were identified in BP, CC, and MF enrichment analysis, respectively (**Figure 3**). The upregulated DEGs significantly enriched in the categories of cell division, mitotic nuclear division, and sister chromatid cohesion, whereas the downregulated DEGs significantly enriched in the categories of epoxide hydrolase P450 pathway, oxidation-reduction processes, and drug metabolic processes, with respect to BP analysis. The upregulated DEGs were mostly involved in the chromosome, centromeric region, midbody, and kinetochore, whereas the downregulated DEGs were mostly involved in the extracellular region, organelle membranes, and extracellular exosomes with respect to CC analysis. The upregulated DEGs mostly enriched in the categories of protein binding, ATP binding, and protein kinase binding, whereas the downregulated DEGs most enriched in the categories of heme binding, oxidoreductase activity, acting on paired donors with incorporation or reduction of molecular oxygen, and oxygen binding with respect to MF analysis.

After KEGG and REACTOME pathway enrichment analyses of DEGs, a total of 51 and 44 enrichment pathways with *P* values <0.05 were identified for upregulated and downregulated DEGs, respectively. The upregulated DEGs mostly enriched in the categories of cell cycle, oocyte meiosis, and progesterone-mediated oocyte maturation, while the downregulated DEGs mostly enriched in the categories of metabolic pathways, retinol metabolism, and tryptophan metabolism with respect to KEGG pathway analysis (**Figure 4**). The upregulated DEGs mostly enriched in the categories of resolution of sister chromatid cohesion, separation of sister chromatids, and mitotic prometaphase, while the downregulated DEGs mostly enriched in the categories of metallothioneins bind-metals, CYP2E1 reactions, and xenobiotics with respect to REACTOME pathway analysis (**Figure 4**).



**Figure 3.** The top 10 functional categories of differentially expressed genes (DEGs) with respect of biological process (BP), cellular component (CC), and molecular function (MF) analysis, respectively. (a) Upregulated DEGs; (b) Downregulated DEGs. GO: 0016705~oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen. 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.

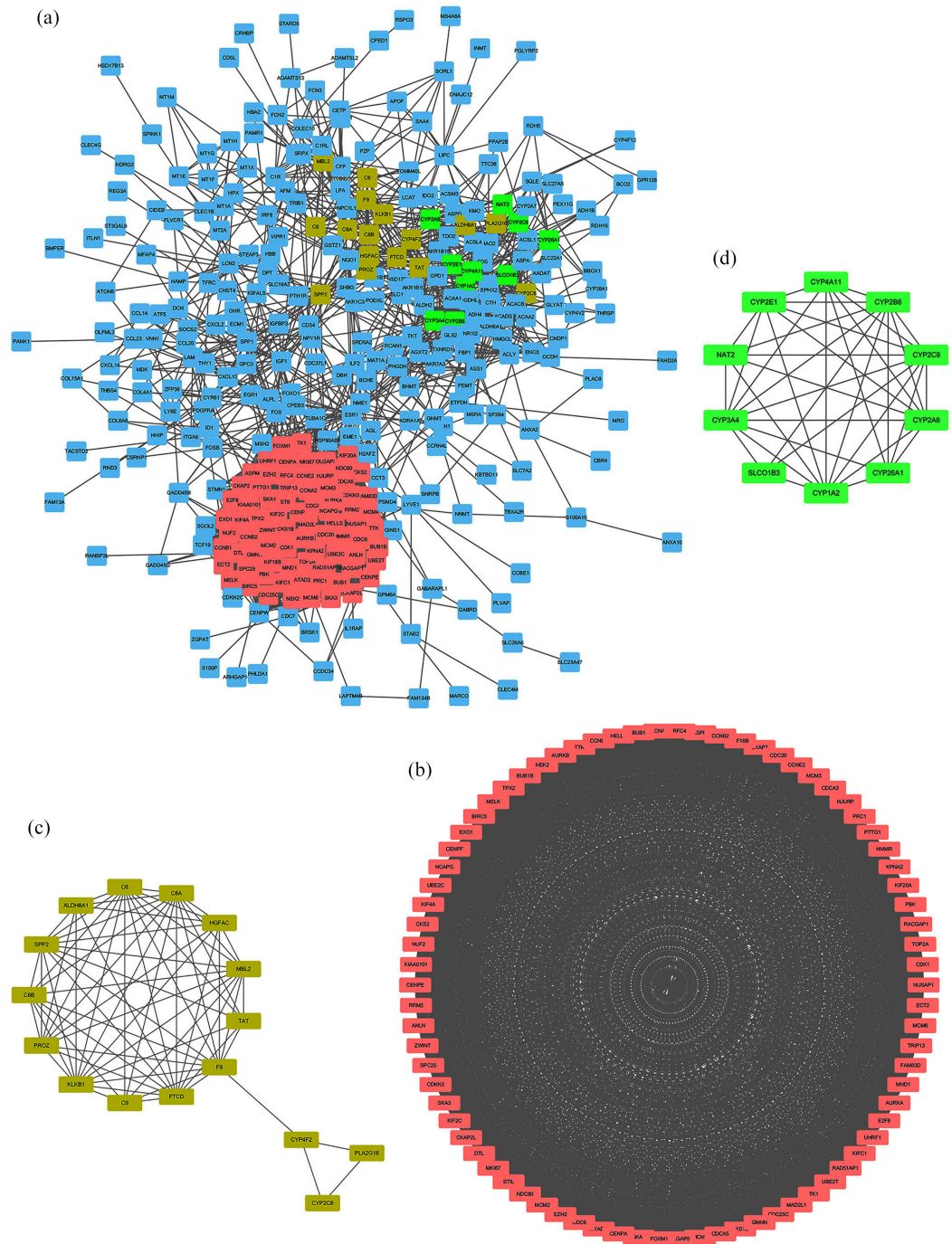


**Figure 4.** KEGG and REACTOME pathway enrichment analysis of differentially expressed genes (DEGs) with P values less than 0.05. (a) Upregulated DEGs; (b) Downregulated DEGs.



### 3.4. Identification of Potential Hub Genes by PPI Network Analysis

A total of 357 DEGs were filtered into the PPI network analysis (Figure 5). By taking the interaction degree of  $\geq 10$ , a total of 165 DEGs that included 104 upregulated DEGs and 61 downregulated DEGs were identified as potential hub genes (Table 3).



**Figure 5.** Protein-protein interaction (PPI) network of differentially expressed genes (DEGs). (a) Total PPI network of DEGs; (b) Module 1; (c) Module 2; (d) Module 3.

**Table 3.** The hub genes identified by protein-protein interaction network and survival analysis.

Upregulated DEGs	Degree	Survival related hub genes	Upregulated DEGs	Degree	Survival related hub genes
CDK1	97	CDK1	UBE2T	67	UBE2T
CCNB1	94	CCNB1	KIFC1 TK1	66	KIFC1 TK1
CCNB2	91	CCNB2	SPC25 EZH2	65	SPC25 EZH2
CDC20 CCNA2 AURKA	90	CDC20 CCNA2 AURKA	CDCA3 HELLS	63	CDCA3 HELLS
MAD2L1 TOP2A	89	MAD2L1 TOP2A	GMNN ATAD2	62	
BUB1B	87	BUB1B	MND1	61	
BUB1 AURKB KIF2C	85	BUB1 AURKB KIF2C	SKA1 KIF18B CCNE2	60	SKA1 KIF18B CCNE2
NDC80 KIF20A TTK RRM2	84	NDC80 KIF20A TTK RRM2	FAM83D E2F8	59	FAM83D E2F8
CDC6 TPX2 DLGAP5 UBE2C		CDC6 TPX2 DLGAP5	CKAP2 CKAP2L	57	CKAP2 CKAP2L
NCAPG BIRC5 NUSAP1 PBK MKI67	83	NCAPG BIRC5 NUSAP1 PBK MKI67	CDC7	53	CDC7
			STIL	50	STIL
ASPM MELK	82	ASPM MELK	GIN51	46	GIN51
CDKN3	81		SGOL2	44	SGOL2
ZWINT NUF2 CENPF		ZWINT NUF2 CENPF	TCF19	36	TCF19
RACGAP1 HMMR PRC1	80	RACGAP1 HMMR PRC1	MSH2 H2AFZ	35	MSH2 H2AFZ
RAD51AP1		RAD51AP1	STMN1 CD34	25	STMN1 CD34
CENPE PTTG1 EXO1 KIF4A		CENPE PTTG1 EXO1 KIF4A	TUBA1C	24	TUBA1C
CDCA5	79	CDCA5	SPP1	23	SPP1
CENPA KIAA0101	78	CENPA KIAA0101	ACLY	21	
FOXM1 DTL	77	FOXM1 DTL	NQO1	20	NQO1
TRIP13	76	TRIP13	PSMD4	19	
NEK2	75	NEK2	EME1	16	EME1
MCM2 MCM4 RFC4	74	MCM2 RFC4	TXNRD1	15	
ECT2	73	ECT2	LCN2	14	
HJURP MCM3 CDC25C	72	HJURP MCM3 CDC25C	CCT3 ITGA6	12	CCT3
MCM6 CKS2 ANLN	70	MCM6 CKS2 ANLN	AKR1C3 TKT THY1 NME1 ILF2	11	AKR1C3 ILF2
KPNA2 UHRF1	69	KPNA2 UHRF1	GPC3 TFRC CDKN2C	10	CDKN2C
SKA3 CKS1B	68	SKA3 CKS1B			
Downregulated DEGs	Degree	Survival related hub genes	Downregulated DEGs	Degree	Survival related hub genes
<i>ESR1</i>	38	<i>ESR1</i>	<i>CXCL12 ACAA1 PCK1</i>	15	<i>PCK1</i>
<i>IGF1 FTCD</i>	26	<i>IGF1 FTCD</i>	<i>C8B ALDH2 LPA CFP NR1 <math>\beta</math> C9</i>	14	<i>C8B ALDH2 LPA CFP NR1 <math>\beta</math> HAO2 ACSL1</i>
<i>CYP3A4</i>	25	<i>CYP3A4</i>	<i>HAO2 ACSL1</i>		
<i>FOS</i>	24		<i>ACADS LCAT CYP2A6 CYP2C8</i>	13	<i>ACADS LCAT CYP2C8</i>
<i>CYP2B6 SPP2</i>	21	<i>SPP2</i>	<i>CYP61 CYP4A11 FOXO1</i>		<i>CYP4A11 FOXO1</i>
<i>C8A CYP2E1 TAT P9</i>	20	<i>C8A CYP2E1 TAT P9</i>	<i>IGFBP3 ACAA2 DCN LIPC</i>	12	<i>LIPC DCN SLC10A1 HGFAC</i>
<i>MBL2 CYP2C9 C6</i>	18	<i>CYP2C9 C6</i>	<i>SLC10A1 C1R HGFAC</i>		
<i>CYP1A2 ASS1</i>	17		<i>KMO ADHA HMGCL CYP4F2</i>	11	<i>ADHA HMGCL CYP4F2</i>
			<i>PLA2G16</i>		
<i>EGR1 CTH CYP26A1 KLKB1</i>			<i>IDO2 ACACB ALPL EPHX2</i>		<i>ALPL EPHX2 ACSMB</i>
<i>ALDH8A1</i>	16	<i>CTH KLKB1 ALDH8A1</i>	<i>ACSM3 GCDH HAMP GNMT</i>	10	<i>GCDH GNMT PROZ</i>
			<i>PROZ ETFDH</i>		<i>ETFDH</i>

DEGs, differentially expressed genes.



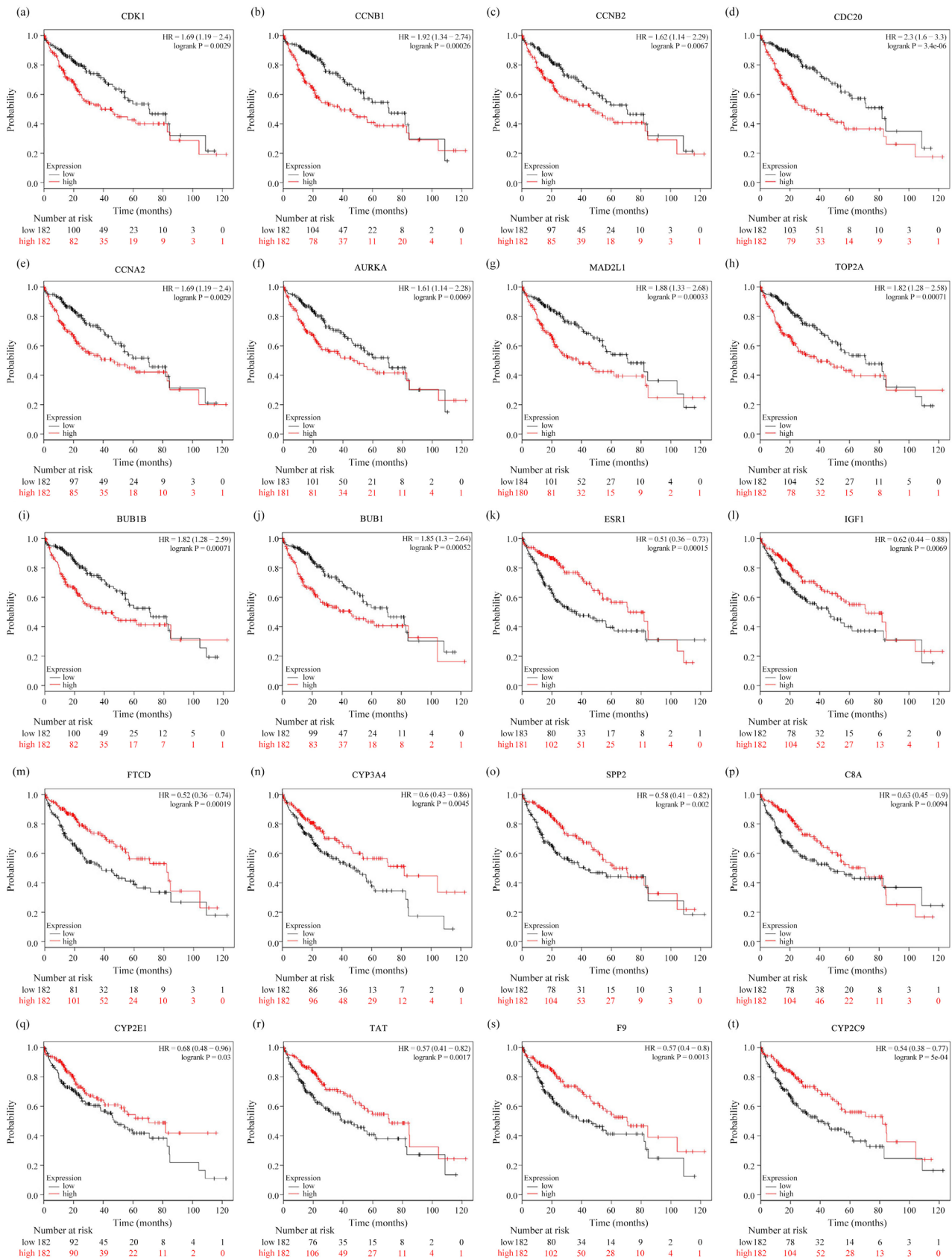
A total of 15 functional modules were found in the MCODE plug-in in Cytoscape, and three significant modules are shown in **Figure 5**. Among these, module 1 included 78 nodules and 2724 edges. The genes in module 1 mainly enriched in the categories of cell division, nucleus, and protein binding with respect to BP, CC, and MF analysis, respectively (**Figure S1**). These genes mostly enriched in the categories of the cell cycle, oocyte meiosis, and progesterone-mediated oocyte maturation with respect to KEGG pathway analysis (**Figure S2**) and enriched in the categories of resolution of sister chromatid cohesion, separation of sister chromatids, and mitotic prometaphase with respect to REACTOME pathway analysis. All DEGs in module 1 were upregulated. Module 2 included 16 nodules and 69 edges. The genes in module 2 mainly enriched in the categories of complement activation, classical pathway, membrane attack complex, serine-type endopeptidase activity, and complement and coagulation cascades pathway after BP, CC, MF and pathway enrichment analysis, respectively (**Figure S1** and **Figure S2**). Module 3 included 10 nodules and 37 edges. The genes in module 3 mainly enriched in the categories of the epoxygenase P450 pathway, organelle membrane, oxygen binding, and xenobiotics pathway with BP, CC, MF and pathway enrichment analysis, respectively (**Figure S1** and **Figure S2**). All DEGs in modules 2 and 3 were downregulated.

### 3.5. Identification of Hub Genes Using Survival Analysis

Because not all potential hub genes are related to HCC prognosis, and thus we identified hub genes that influence HCC survival. After a total of 88 upregulated genes and 40 downregulated genes were identified as hub genes. The top 10 upregulated hub genes were *CDK1*, *CCNB1*, *CCNB2*, *CDC20*, *CCNA2*, *AURKA*, *MAD2L1*, *TOP2A*, *BUB1B*, and *BUB1*. The top 10 downregulated hub genes were *ESR1*, *IGF1*, *FTCD*, *CYP3A4*, *SPP2*, *C8A*, *CYP2E1*, *TAT*, *F9*, and *CYP2C9*. The results of survival analysis of the top 10 upregulated and top 10 downregulated hub genes are shown in **Figure 6**.

## 4. Discussion

The identification of hub genes by bioinformatics analysis is widely used in tumorigenesis research, including HCC [9] [10]. However, most studies employ the Venn diagram to identify hub genes [9] [10], and thus some important DEGs may be missed. In addition, not all research studies use the same gene chip platform and not all gene chip platforms have all the gene probes. In addition, over time, more genes are being discovered, so that some genes that have not previously been detected by microarrays may be differentially expressed genes and have important functions. Robust rank aggregation algorithms can avoid these discrepancies and integrate data from different platforms [11]. Another difference from previous studies is that the minimum sample size of the dataset included in this study is 39, whereas that of previous studies is <10 [9] [10]. The small sample size of datasets may lead to biased or statistically insignificant results. This study identified more hub genes than previous studies because this



**Figure 6.** Overall survival analysis of the top 10 upregulated and downregulated hub genes.

study utilized datasets with a larger sample size and integrated DEG data using an RRA algorithm.

Our study identified 151 upregulated DEGs, which is relatively higher in number than previous studies [9] [10]. Through GO analysis, these genes were determined to be mostly enriched in the functional categories of cell division, chromosome, centromeric region, and protein binding using BP, CC and MF analysis, respectively. Pathway analysis revealed that these genes enriched in the categories of the cell cycle, resolution of sister chromatid cohesion, and separation of sister chromatids pathway. These results are similar to those of previous studies [10] [12]. Module 1 is the most significant network complex that was filtered by PPI network analysis, and all genes are upregulated. The DEGs in Module 1 is similar to that described by Shen [10]. These genes play an important role in tumor cell cycle. Blocking the expression of any gene may lead to a disruption of the cell cycle, thereby delaying the occurrence and development of tumors. In addition, we also identified a total of 245 downregulated DEGs. Through GO analysis, these genes were shown to be mostly enriched in the categories of epoxigenase P450 pathway, extracellular region, and heme binding with BP, CC, and MF analysis, respectively. Pathway analysis showed that these genes enriched in metabolic pathways, metal-binding metallothioneins bind-metals, and retinol metabolism. These results have also been described in previous studies [10] [13]. Module 2 is the second most significant network complex that was filtered using PPI network analysis, and the genes in module 2 are all downregulated. However, the genes in module 2 differ from that described by Shen [10]. Furthermore, the genes in module 2 enriched in the complement and coagulation cascades pathway. The genes in module 3 enriched in the xenobiotics pathway, which is similar to module 2 of Shen's study [10], but the genes in module 3 in our study significantly vary from that in module 2 of the same study [10]. These discrepancies may be attributed to the sample size of the datasets employed in the analysis and variations in the integration algorithm employed [10].

PPI network analysis was used to identify potential hub genes. Because not all DEGs are associated with HCC prognosis, we identified hub genes that may be related to HCC prognosis. We identified 141 hub genes, which is relatively higher than that reported by previous studies [9] [10] [12] [13]. Most of the DEGs with high interactive degree were upregulated. The top 10 upregulated DEGs were *CDK1*, *CCNB1*, *CCNB2*, *CDC20*, *CCNA2*, *AURKA*, *MAD2L1*, *TOP2A*, *BUB1B*, and *BUB1*. These results partially differ from the findings of previous studies [9] [10] [12] [13].

The *CDK1* gene showed the highest degree of interaction in the PPI network. *CDK1* is an M-phase promoting factor that plays a key role in eukaryotic cell cycle by modulating G1/S and G2/M phase transitions. Several studies have elucidated the function of *CDK1* in HCC [14] [15] [16]. Zhang *et al.* reported that miR-582-5p can inhibit cell proliferation and induce cell cycle arrest at the G0/G1 phase of HCC by targeting *CDK1* and *AKT3* [15]. Zhou *et al.* reported that metformin can induce miR-378 to downregulate *CDK1*, which in turn inhi-

bits cell proliferation in HCC via G2/M arrest [16]. Wu *et al.* reported that the *CDK1* inhibitor RO3306 can improve the efficacy of sorafenib in the treatment of HCC patient-derived xenograft tumor models by blocking the *CDK1/PDK1/β-catenin* signaling pathway [14]. *CCNB1* and *CCNB2* encode proteins that are associated with *CDK1* and are essential to the cell cycle by regulating the G2/M transition. *CCNA2* controls G1/S and G2/M phase transitions by forming a serine/threonine protein kinase complex with *CDK1* or *CDK2*. Gu *et al.* reported that microRNA-144 can inhibit cell proliferation, migration, and invasion of HCC by targeting *CCNB1* [17]. Chai *et al.* reported that *FOXM1* can promote cell proliferation in HCC by directly binding and activating the *CCNB1* gene at the transcriptional level [18]. Li *et al.* reported that the overexpression of *CCNB2* promotes cell proliferation and migration of HCC via the *CCNB2/PLK1* pathway, and *CCNB2* is associated with poor prognosis [19]. Yue *et al.* reported that knocking down *CCNA2* reduces cell proliferation, which is caused by *ZHX2* knockdown [20]. *CDC20* act as a regulatory protein at multiple points in the cell cycle by interacting with other proteins. *CDC20* is required for the activity of the anaphase-promoting complex/cyclosome, and the complex is modulated by *MAD2L1*, which is a component of the mitotic spindle assembly checkpoint. Li *et al.* reported that silencing *CDC20* can inhibit cell proliferation by G2/M-phase arrest [21]. Li *et al.* and colleagues reported that miR-200c-5p can suppress cell proliferation, migration, and invasion of HCC by downregulating *MAD2L1* [22]. *AURKA* is a mitotic serine/threonine kinase that regulates cell cycle progression. Chen *et al.* reported that the overexpression of *AURKA* induces entry into the epithelial-mesenchymal transition and cancer stem cell behaviors via the PI3K/AKT pathway and silences *AURKA*-suppressed radiation-enhanced cell invasiveness in HCC [23]. Gao *et al.* reported that downregulation of *AURKA* results in G2/M phase cell arrest and induces apoptosis of HepG2 cells [24]. Also, Zhang *et al.* reported that *AURKA* promotes chemoresistance in HCC by targeting the NF-kappaB/microRNA-21/PTEN signaling pathway [25]. *TOP2A* is essential for proper segregation of daughter chromosomes during mitosis and meiosis by controlling the topological states of DNA. Sudan *et al.* reported that quercetin-3-O-glucoside induces cell cycle arrest in HCC by suppressing *TOP2A* [26]. Both *BUB1B* and *BUB1* encode serine/threonine-protein kinases that are involved in spindle checkpoint functions during mitosis [27]. Xu *et al.* reported that MiR-490-5p could inhibit cell proliferation, invasion, and migration as well as increase apoptosis by regulating the TGFβ/Smad signaling pathways by inhibiting *BUB1* [28]. The top 10 upregulated DEGs are involved in mitosis and cell cycle regulation.

This study also identified several downregulated hub genes that enriched in various metabolic pathways. However, these genes are also associated with HCC prognosis, indicating that these genes play an important role in HCC proliferation and progression. The top 10 downregulated DEGs were *ESR1*, *IGF1*, *FTCD*, *CYP3A4*, *SPP2*, *C8A*, *CYP2E1*, *TAT*, *P9*, and *CYP2C9*. *ESR1* encodes an estrogen receptor that binds with a ligand steroid hormone to regulate gene expression and cell proliferation and differentiation in eukaryotic tissues. Tu *et al.* re-

ported that the overexpression of *ESR1* induces apoptosis in *ESR1*-negative Hep3B cells by binding *ESR1* to *SP1* protein [29]. Additionally, this *ESR1*-*SP1* complex binds to the *TNF $\alpha$*  gene promoter and active caspase 3 in a ligand-dependent manner [29]. Liu *et al.* reported that microRNA-18a stimulates cell proliferation in HCC by suppressing *ESR1* expression [30]. Therefore, microRNA-18a could block the protective effects of estrogen and promote the development of HCC in women [30]. Additionally, Dai *et al.* reported that *ESR1* may be a tumor suppressor gene in HCC, and *ESR1* may be repressed by promoter hypermethylation [31]. *IGF1* encodes a protein that is similar to insulin and is involved in mediating growth and development. This study has also shown that *IGF1* is positively correlated with HCC prognosis. In addition, Shen *et al.* reported that preoperative low serum levels of *IGF-1* indicate poor prognosis in HCC patients who have undergone hepatectomy [32], and Cho *et al.* reported that low baseline serum *IGF-1* levels are independently correlated with shorter time to progression and poorer overall survival in patients with HCC who underwent TACE [33]. Conversely, several studies have found that the *IGF1* is positively correlated with HCC tumorigenesis [34] [35]. Liu reported that *IGF-1* promotes the invasive and migratory ability of HCC cells by epithelial-mesenchymal transition via activating survivin [35]. Lei *et al.* reported that *IGF-1* induces the growth and metastasis of HCC by inhibiting proteasome-mediated *CTSB* degradation [34]. Therefore, the function and mechanism of *IGF1* in HCC tumorigenesis still need to be verified in future studies. *FTCD* encodes a protein that participates in histidine catabolism. Naama and colleagues reported that histidine catabolism drains the cellular pool of tetrahydrofolate, which is an essential cofactor in nucleotide synthesis [36]. Thus, upregulating *FTCD* may help inhibit cell proliferation. *CYP3A4*, *CYP2E1*, and *CYP2C9* encode several members of the cytochrome P450 superfamily of enzymes. Cytochrome P450 proteins can catalyze many reactions that are involved in the synthesis of cholesterol and steroids and in drug metabolism. Ashida *et al.* identified that *CYP3A4* is a novel tumor suppressor gene, and its downregulation is related to poor HCC prognosis [37]. Liu *et al.* reported that *HBx* promotes cell growth by inhibiting *CYP2E1* expression by downregulating HNF4 $\alpha$  [38]. Yu *et al.* reported that *CYP2C9* is suppressed in HCC cells by hsa-miR-128-3p [39]. These studies indicate that *CYP2C9*, *CYP2E1*, and *CYP3A4* are protective biomarkers for HCC patients, but their underlying mechanisms remain unclear [37] [38] [39]. These possibly inhibit tumorigenesis by regulating metabolic pathways [40] [41], as indicated by the results of pathway enrichment analysis in this study. *C8A* encodes protein that is a component of the complement cascade system. *F9* encodes a protein that is involved in the blood coagulation cascade. However, these two genes also positively correlated to overall survival in HCC, although the underlying mechanisms are unknown. These may involve regulation of the complement and coagulation cascades pathway in which the two genes are involved, as demonstrated in the results of KEGG pathway enrichment analysis. Lao *et al.* reported that phosphoprotein *SPP2* can inhibit the growth and bone metastasis of *BMP2*-

induced prostate cancer [42]. However, the mechanism of lower *SPP2* expression that is associated with poor prognosis of HCC remains unclear. Fu and colleagues reported that the downregulation of *TAT* contributes to the development and progression of HCC [43]. However, the mechanism of *TAT* in the tumorigenesis inhibition of HCC is not clear. These genes may also become novel targets for the treatment of HCC, and the mechanisms of these genes may be elucidated in future studies.

The hub genes identified in this study are much more reliable than hub genes identified in previous studies although this study identified much more hub DEGs than previous studies. This may be due to the use of a larger sample size for the included datasets, thereby reducing selection bias. The second reason is that the RRA algorithm in integrating DEGs may decrease data omissions. The third reason is that the DEGs were associated with the prognosis of HCC. However, the study also has several limitations. First, our prognostic analysis is based on TCGA data only. The sample size is relatively small and there is a certain bias risk. Second, these DEGs still need to be validated by *in vivo* and *in vitro* studies, as there were only identified by bioinformatics analysis. Also, the mechanism of several DEGs still need to be elucidated.

## 5. Conclusion

This study identified a significantly higher number of DEGs that are associated with HCC prognosis. Upregulated DEGs enriched in the cell cycle, whereas down-regulated DEGs enriched in the metabolism pathway. These results needed to be verified by *in vivo* and *in vitro* studies.

## Data Availability

The gene expression data supporting this bioinformatics analysis are from the Gene Expression Omnibus (<https://www.ncbi.nlm.nih.gov/geo>) and The Cancer Genome Atlas liver hepatocellular carcinoma (<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>) datasets.

## Acknowledgements

This study was funding from the Technical Innovation and Application Development Project of Chongqing Science and Technology Bureau (cstc2019jcxmsxmX0144), Natural Science Foundation of Chongqing (No. cstc2018jcyjAX0800).

## Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

## References

- [1] Siegel, R.L., Miller, K.D. and Jemal, A. (2017) Cancer Statistics, 2017. *CA: A Cancer*



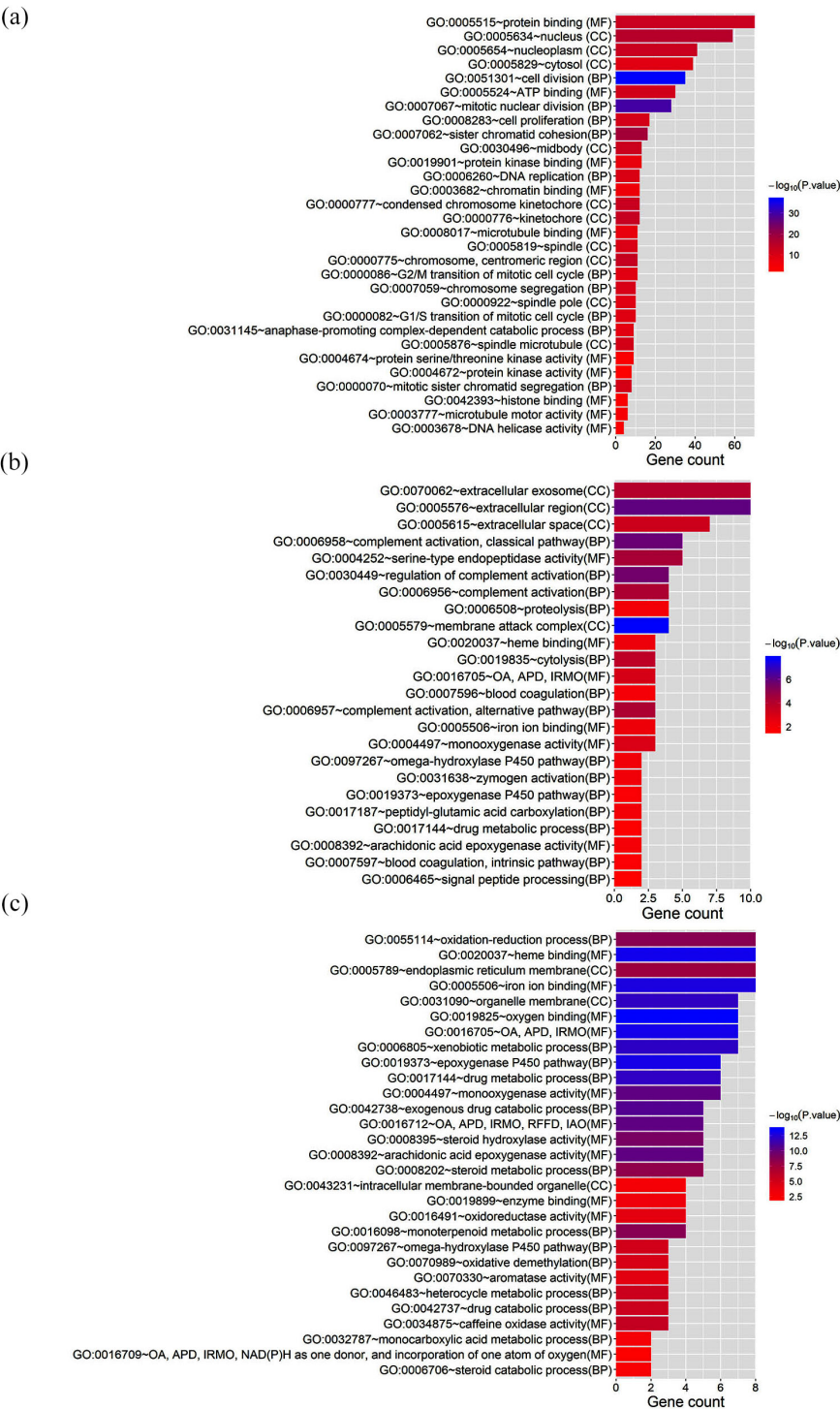
- Journal for Clinicians*, **67**, 7-30. <https://doi.org/10.3322/caac.21387>
- [2] Forner, A., Reig, M. and Bruix, J. (2018) Hepatocellular Carcinoma. *The Lancet*, **391**, 1301-1314. [https://doi.org/10.1016/S0140-6736\(18\)30010-2](https://doi.org/10.1016/S0140-6736(18)30010-2)
  - [3] Huang, G., Li, P.P., Lau, W.Y., Pan, Z.Y., Zhao, L.H., Wang, Z.G., *et al.* (2018) Antiviral Therapy Reduces Hepatocellular Carcinoma Recurrence in Patients with Low HBV-DNA Levels: A Randomized Controlled Trial. *Annals of Surgery*, **268**, 943-954. <https://doi.org/10.1097/SLA.0000000000002727>
  - [4] Llovet, J.M., Ricci, S., Mazzaferro, V., Hilgard, P., Gane, E., Blanc, J.F., *et al.* (2008) Sorafenib in Advanced Hepatocellular Carcinoma. *The New England Journal of Medicine*, **359**, 378-390. <https://doi.org/10.1056/NEJMoa0708857>
  - [5] Bruix, J., Takayama, T., Mazzaferro, V., Chau, G.Y., Yang, J., Kudo, M., *et al.* (2015) Adjuvant Sorafenib for Hepatocellular Carcinoma after Resection or Ablation (STORM): A Phase 3, Randomised, Double-Blind, Placebo-Controlled Trial. *The Lancet Oncology*, **16**, 1344-1354. [https://doi.org/10.1016/S1470-2045\(15\)00198-9](https://doi.org/10.1016/S1470-2045(15)00198-9)
  - [6] Lim, H.Y., Sohn, I., Deng, S., Lee, J., Jung, S.H., Mao, M., *et al.* (2013) Prediction of Disease-Free Survival in Hepatocellular Carcinoma by Gene Expression Profiling. *Annals of Surgical Oncology*, **20**, 3747-3753. <https://doi.org/10.1245/s10434-013-3070-y>
  - [7] Roessler, S., Jia, H.L., Budhu, A., Forgues, M., Ye, Q.H., Lee, J.S., *et al.* (2010) A Unique Metastasis Gene Signature Enables Prediction of Tumor Relapse in Early-Stage Hepatocellular Carcinoma Patients. *Cancer Research*, **70**, 10202-10212. <https://doi.org/10.1158/0008-5472.CAN-10-2607>
  - [8] Tung, E.K., Mak, C.K., Fatima, S., Lo, R.C., Zhao, H., Zhang, C., *et al.* (2011) Clinicopathological and Prognostic Significance of Serum and Tissue Dickkopf-1 Levels in Human Hepatocellular Carcinoma. *Liver International*, **31**, 1494-1504. <https://doi.org/10.1111/j.1478-3231.2011.02597.x>
  - [9] Chen, Z., Chen, J., Huang, X., Wu, Y., Huang, K., Xu, W., *et al.* (2019) Identification of Potential Key Genes for Hepatitis B Virus-Associated Hepatocellular Carcinoma by Bioinformatics Analysis. *Journal of Computational Biology*, **26**, 485-494. <https://doi.org/10.1089/cmb.2018.0244>
  - [10] Shen, S., Kong, J., Qiu, Y., Yang, X., Wang, W. and Yan, L. (2019) Identification of Core Genes and Outcomes in Hepatocellular Carcinoma by Bioinformatics Analysis. *Journal of Cellular Biochemistry*, **120**, 10069-10081. <https://doi.org/10.1002/jcb.28290>
  - [11] Kolde, R., Laur, S., Adler, P. and Vilo, J. (2012) Robust Rank Aggregation for Gene List Integration and Meta-Analysis. *Bioinformatics*, **28**, 573-580. <https://doi.org/10.1093/bioinformatics/btr709>
  - [12] Sun, B., Lin, G., Ji, D., Li, S., Chi, G. and Jin, X. (2018) Dysfunction of Sister Chromatids Separation Promotes Progression of Hepatocellular Carcinoma According to Analysis of Gene Expression Profiling. *Frontiers in Physiology*, **9**, 1019. <https://doi.org/10.3389/fphys.2018.01019>
  - [13] Gao, X., Wang, X. and Zhang, S. (2018) Bioinformatics Identification of Crucial Genes and Pathways Associated with Hepatocellular Carcinoma. *Bioscience Reports*, **38**, BSR20181441. <https://doi.org/10.1042/BSR20181441>
  - [14] Wu, C.X., Wang, X.Q., Chok, S.H., Man, K., Tsang, S.H.Y., Chan, A.C.Y., *et al.* (2018) Blocking CDK1/PDK1/beta-Catenin Signaling by CDK1 Inhibitor RO3306 Increased the Efficacy of Sorafenib Treatment by Targeting Cancer Stem Cells in a Preclinical Model of Hepatocellular Carcinoma. *Theranostics*, **8**, 3737-3750.

- <https://doi.org/10.7150/thno.25487>
- [15] Zhang, Y., Huang, W., Ran, Y., Xiong, Y., Zhong, Z., Fan, X., et al. (2015) miR-582-5p Inhibits Proliferation of Hepatocellular Carcinoma by Targeting CDK1 and AKT3. *Tumor Biology*, **36**, 8309-8316. <https://doi.org/10.1007/s13277-015-3582-0>
  - [16] Zhou, J., Han, S., Qian, W., Gu, Y., Li, X. and Yang, K. (2018) Metformin Induces miR-378 to Downregulate the CDK1, Leading to Suppression of Cell Proliferation in Hepatocellular Carcinoma. *Oncotargets and Therapy*, **11**, 4451-4459. <https://doi.org/10.2147/OTT.S167614>
  - [17] Gu, J., Liu, X., Li, J. and He, Y. (2019) MicroRNA-144 Inhibits Cell Proliferation, Migration and Invasion in Human Hepatocellular Carcinoma by Targeting CCNB1. *Cancer Cell International*, **19**, 15. <https://doi.org/10.1186/s12935-019-0729-x>
  - [18] Chai, N., Xie, H.H., Yin, J.P., Sa, K.D., Guo, Y., Wang, M., et al. (2018) FOXM1 Promotes Proliferation in Human Hepatocellular Carcinoma Cells by Transcriptional Activation of CCNB1. *Biochemical and Biophysical Research Communications*, **500**, 924-929. <https://doi.org/10.1016/j.bbrc.2018.04.201>
  - [19] Li, R., Jiang, X., Zhang, Y., Wang, S., Chen, X., Yu, X., et al. (2019) Cyclin B2 Overexpression in Human Hepatocellular Carcinoma Is Associated with Poor Prognosis. *Archives of Medical Research*, **50**, 10-17. <https://doi.org/10.1016/j.arcmed.2019.03.003>
  - [20] Yue, X., Zhang, Z., Liang, X., Gao, L., Zhang, X., Zhao, D., et al. (2012) Zinc Fingers and Homeobox 2 Inhibits Hepatocellular Carcinoma Cell Proliferation and Represses Expression of Cyclins A and E. *Gastroenterology*, **142**, 1559-70e2. <https://doi.org/10.1053/j.gastro.2012.02.049>
  - [21] Li, J., Gao, J.Z., Du, J.L., Huang, Z.X. and Wei, L.X. (2014) Increased CDC20 Expression Is Associated with Development and Progression of Hepatocellular Carcinoma. *International Journal of Oncology*, **45**, 1547-1555. <https://doi.org/10.3892/ijo.2014.2559>
  - [22] Li, Y., Bai, W. and Zhang, J. (2017) MiR-200c-5p Suppresses Proliferation and Metastasis of Human Hepatocellular Carcinoma (HCC) via Suppressing MAD2L1. *Biomedicine & Pharmacotherapy*, **92**, 1038-1044. <https://doi.org/10.1016/j.biopha.2017.05.092>
  - [23] Chen, C., Song, G., Xiang, J., Zhang, H., Zhao, S. and Zhan, Y. (2017) AURKA Promotes Cancer Metastasis by Regulating Epithelial-Mesenchymal Transition and Cancer Stem Cell Properties in Hepatocellular Carcinoma. *Biochemical and Biophysical Research Communications*, **486**, 514-520. <https://doi.org/10.1016/j.bbrc.2017.03.075>
  - [24] Gao, P., Wang, R., Shen, J.J., Lin, F., Wang, X., Dong, K., et al. (2008) Hypoxia-Inducible Enhancer/Alpha-Fetoprotein Promoter-Driven RNA Interference Targeting STK15 Suppresses Proliferation and Induces Apoptosis in Human Hepatocellular Carcinoma Cells. *Cancer Science*, **99**, 2209-2217. <https://doi.org/10.1111/j.1349-7006.2008.00941.x>
  - [25] Zhang, K., Chen, J., Chen, D., Huang, J., Feng, B., Han, S., et al. (2014) Aurora-A Promotes Chemoresistance in Hepatocellular Carcinoma by Targeting NF-kappaB/microRNA-21/PTEN Signaling Pathway. *Oncotarget*, **5**, 12916-12935. <https://doi.org/10.18632/oncotarget.2682>
  - [26] Sudan, S. and Rupasinghe, H.P. (2014) Quercetin-3-O-Glucoside Induces Human DNA Topoisomerase II Inhibition, Cell Cycle Arrest and Apoptosis in Hepatocellular Carcinoma Cells. *Anticancer Research*, **34**, 1691-1699.
  - [27] Raaijmakers, J.A., van Heesbeen, R., Blomen, V.A., Janssen, L.M.E., van Diemen,

- F., Brummelkamp, T.R., *et al.* (2018) BUB1 Is Essential for the Viability of Human Cells in Which the Spindle Assembly Checkpoint Is Compromised. *Cell Reports*, **22**, 1424-1438. <https://doi.org/10.1016/j.celrep.2018.01.034>
- [28] Xu, B., Xu, T., Liu, H., Min, Q., Wang, S. and Song, Q. (2017) MiR-490-5p Suppresses Cell Proliferation and Invasion by Targeting BUB1 in Hepatocellular Carcinoma Cells. *Pharmacology*, **100**, 269-282. <https://doi.org/10.1159/000477667>
- [29] Tu, C.C., Kumar, V.B., Day, C.H., Kuo, W.W., Yeh, S.P., Chen, R.J., *et al.* (2013) Estrogen Receptor Alpha (ESR1) Over-Expression Mediated Apoptosis in Hep3B Cells by Binding with SP1 Proteins. *Journal of Molecular Endocrinology*, **51**, 203-212. <https://doi.org/10.1530/JME-13-0085>
- [30] Liu, W.H., Yeh, S.H., Lu, C.C., Yu, S.L., Chen, H.Y., Lin, C.Y., *et al.* (2009) Micro-RNA-18a Prevents Estrogen Receptor-Alpha Expression, Promoting Proliferation of Hepatocellular Carcinoma cells. *Gastroenterology*, **136**, 683-693. <https://doi.org/10.1053/j.gastro.2008.10.029>
- [31] Dai, B., Geng, L., Yu, Y., Sui, C., Xie, F., Shen, W., *et al.* (2014) Methylation Patterns of Estrogen Receptor Alpha Promoter Correlate with Estrogen Receptor Alpha Expression and Clinicopathological Factors in Hepatocellular Carcinoma. *Experimental Biology and Medicine (Maywood)*, **239**, 883-890. <https://doi.org/10.1177/1535370214536651>
- [32] Shen, L., Xu, L., Zhang, J. and Jiang, D. (2018) Preoperative Serum Insulin-Like Growth Factor 1 Level as a Prognostic Factor in Patients Undergoing Hepatic Resection for Hepatocellular Carcinoma. *Journal of Interferon & Cytokine Research*, **38**, 153-160. <https://doi.org/10.1089/jir.2017.0107>
- [33] Cho, E., Kim, H.C., Lee, J.H., Jeong-Ju, Y., Choi, W.M., Cho, Y.Y., *et al.* (2014) Serum Insulin-Like Growth Factor-1 Predicts Disease Progression and Survival in Patients with Hepatocellular Carcinoma Who Undergo Transarterial Chemoembolization. *PLoS ONE*, **9**, e90862. <https://doi.org/10.1371/journal.pone.0090862>
- [34] Lei, T. and Ling, X. (2015) IGF-1 Promotes the Growth and Metastasis of Hepatocellular Carcinoma via the Inhibition of Proteasome-Mediated Cathepsin B Degradation. *World Journal of Gastroenterology*, **21**, 10137-10149. <https://doi.org/10.3748/wjg.v21.i35.10137>
- [35] Liu, F., Sun, Y., Liu, B., Lu, J., Li, H., Zhu, H., *et al.* (2018) Insulin-Like Growth Factor-1 Induces Epithelial-Mesenchymal Transition in Hepatocellular Carcinoma by Activating Survivin. *Oncology Reports*, **40**, 952-958. <https://doi.org/10.3892/or.2018.6516>
- [36] Kanarek, N., Keys, H.R., Cantor, J.R., Lewis, C.A., Chan, S.H., Kunchok, T., *et al.* (2018) Histidine Catabolism Is a Major Determinant of Methotrexate Sensitivity. *Nature*, **559**, 632-636. <https://doi.org/10.1038/s41586-018-0316-7>
- [37] Ashida, R., Okamura, Y., Ohshima, K., Kakuda, Y., Uesaka, K., Sugiura, T., *et al.* (2017) CYP3A4 Gene Is a Novel Biomarker for Predicting a Poor Prognosis in Hepatocellular Carcinoma. *Cancer Genomics Proteomics*, **14**, 445-453. <https://doi.org/10.21873/cgp.20054>
- [38] Liu, H., Lou, G., Li, C., Wang, X., Cederbaum, A.I., Gan, L., *et al.* (2014) HBx Inhibits CYP2E1 Gene Expression via Downregulating HNF4alpha in Human Hepatoma Cells. *PLoS ONE*, **9**, e107913. <https://doi.org/10.1371/journal.pone.0107913>
- [39] Yu, D., Green, B., Marrone, A., Guo, Y., Kadlubar, S., Lin, D., *et al.* (2015) Suppression of CYP2C9 by microRNA hsa-miR-128-3p in Human Liver Cells and Association with Hepatocellular Carcinoma. *Scientific Reports*, **5**, Article No. 8534. <https://doi.org/10.1038/srep08534>

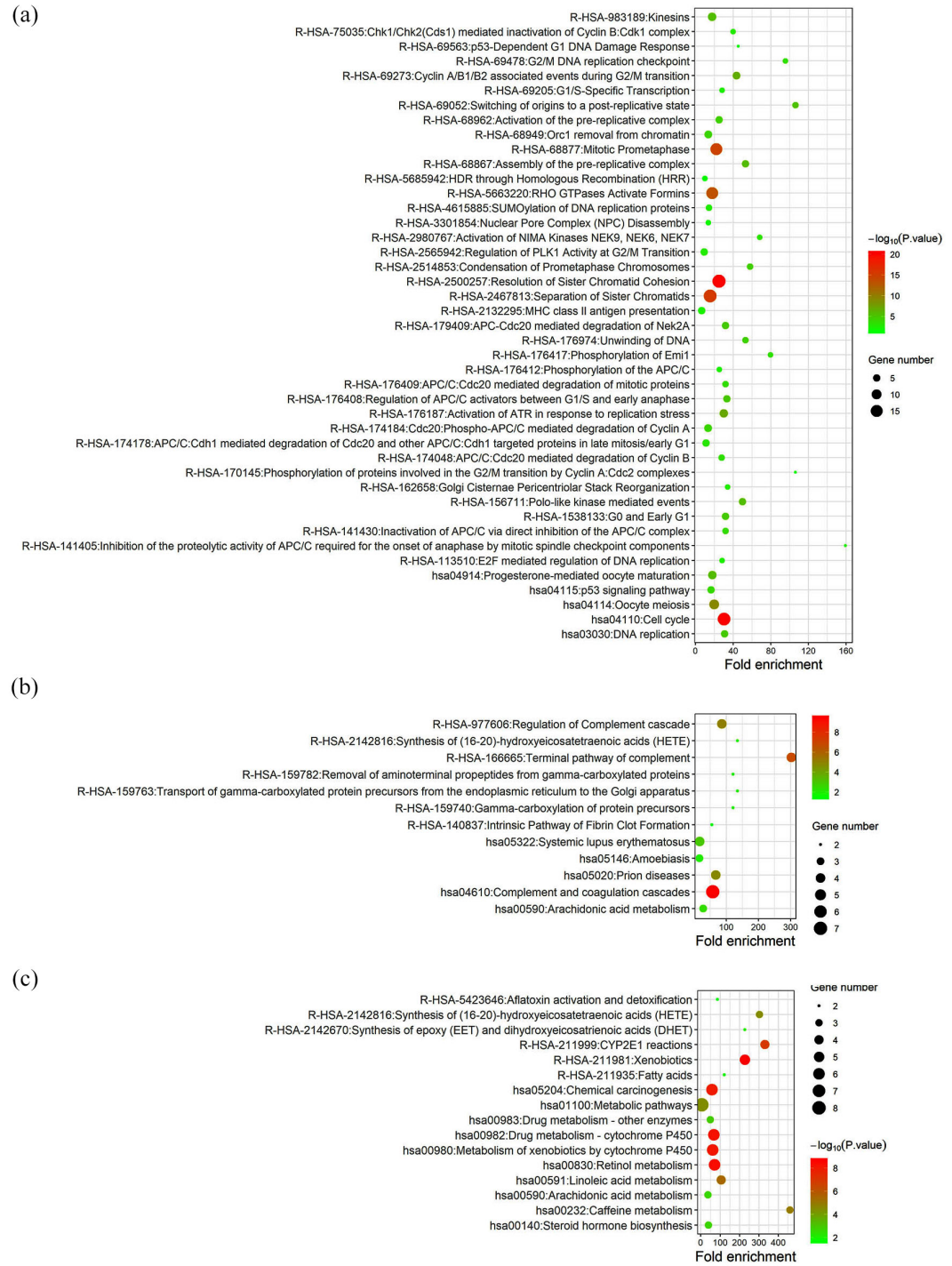
- [40] Tsunedomi, R., Iizuka, N., Hamamoto, Y., Uchimura, S., Miyamoto, T., Tamesa, T., *et al.* (2005) Patterns of Expression of Cytochrome P450 Genes in Progression of hepatitis C Virus-Associated Hepatocellular Carcinoma. *International Journal of Oncology*, **27**, 661-667
- [41] Yan, T., Lu, L., Xie, C., Chen, J., Peng, X., Zhu, L., *et al.* (2015) Severely Impaired and Dysregulated Cytochrome P450 Expression and Activities in Hepatocellular Carcinoma: Implications for Personalized Treatment in Patients. *Molecular Cancer Therapeutics*, **14**, 2874-2886. <https://doi.org/10.1158/1535-7163.MCT-15-0274>
- [42] Lao, L., Shen, J., Tian, H., Yao, Q., Li, Y., Qian, L., *et al.* (2016) Secreted Phosphoprotein 24 kD Inhibits Growth of Human Prostate Cancer Cells Stimulated by BMP-2. *Anticancer Research*, **36**, 5773-5780. <https://doi.org/10.21873/anticancerres.11161>
- [43] Fu, L., Dong, S.S., Xie, Y.W., Tai, L.S., Chen, L., Kong, K.L., *et al.* (2010) Down-Regulation of Tyrosine Aminotransferase at a Frequently Deleted Region 16q22 Contributes to the Pathogenesis of Hepatocellular Carcinoma. *Hepatology*, **51**, 1624-1634. <https://doi.org/10.1002/hep.23540>

Supplementary Materials



**Figure S1.** The top 10 functional categories of genes in module 1 (a), 2 (b) and 3 (c) with respect to biological process, cellular component and molecular function analysis.





**Figure S2.** The KEGG and REACTOME pathway enrichment analysis of genes in module 1 (a), 2 (b) and 3 (c).