

Research Article

Identification of Potential Hub Genes Related to Diagnosis and Prognosis of Hepatitis B Virus-Related Hepatocellular Carcinoma via Integrated Bioinformatics Analysis

Yuqin Tang,^{1,2} Yongqiang Zhang^{3,4} and Xun Hu²

¹School of Basic Medical Sciences, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, China

²Biorepository, State Key Laboratory of Biotherapy and Cancer Center, West China Hospital, Sichuan University, Chengdu 610041, China

³Molecular Medicine Center, West China Hospital, Sichuan University, Chengdu 610041, China

⁴West China School of Medicine, West China Hospital, Sichuan University, Chengdu 610041, China

Correspondence should be addressed to Xun Hu; hxxhu99@163.com

Received 8 May 2020; Revised 4 August 2020; Accepted 13 August 2020; Published 9 December 2020

Academic Editor: Nagarajan Raju

Copyright © 2020 Yuqin Tang et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Hepatocellular carcinoma (HCC) is a common malignant cancer with poor survival outcomes, and hepatitis B virus (HBV) infection is most likely to contribute to HCC. But the molecular mechanism remains obscure. Our study intended to identify the candidate potential hub genes associated with the carcinogenesis of HBV-related HCC (HBV-HCC), which may be helpful in developing novel tumor biomarkers for potential targeted therapies. Four transcriptome datasets (GSE84402, GSE25097, GSE94660, and GSE121248) were used to screen the 309 overlapping differentially expressed genes (DEGs), including 100 upregulated genes and 209 downregulated genes. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment were used to explore the biological function of DEGs. A PPI network based on the STRING database was constructed and visualized by the Cytoscape software, consisting of 209 nodes and 1676 edges. Then, we recognized 17 hub genes by CytoHubba plugin, which were further validated on additional three datasets (GSE14520, TCGA-LIHC, and ICGC-LIRI-JP). The diagnostic effectiveness of hub genes was assessed with receiver operating characteristic (ROC) analysis, and all hub genes displayed good performance in discriminating TNM stage I patient samples and normal tissue ones. For prognostic analysis, two prognostic key genes (TOP2A and KIF11) out of the 17 hub genes were screened and used to develop a prognostic signature, which showed good potential for overall survival (OS) stratification of HBV-HCC patients. Gene Set Enrichment Analysis (GSEA) was performed in order to better understand the function of this prognostic gene signature. Finally, the miRNA-mRNA regulatory relationships of all hub genes in human liver were predicted using miRNet. In conclusion, the current study gives further insight on the pathogenesis and carcinogenesis of HBV-HCC, and the identified DEGs provide a promising direction for improving the diagnostic, prognostic, and therapeutic outcomes of HBV-HCC.

1. Introduction

Liver cancer, with about 841,000 new cases diagnosed and 782,000 deaths in 2018, still represents a common lethal solid tumor and ranks fourth leading cause of cancer-related deaths worldwide [1]. In China, liver cancer was one of the first five life killers in 2017 [2]. Hepatocellular carcinoma (HCC), comprising 75%-85% of all primary liver cancer cases worldwide [1], is the primary histological subtype. The major

causative etiological factors of HCC are considered as infection of endemic hepatitis B virus (HBV) or hepatitis C virus (HCV), followed by exposition to aflatoxin B1, alcohol abuse, and obesity [1, 3]. Particularly, HBV infection is considered the dominant cause of HCC, accounting for more than 80% of all HCC incidences in China and other developing countries [4]. Despite significant advances in early diagnosis, prevention and the standard therapeutic interventions such as surgery, radiation, chemotherapy, or personalized target

therapeutic strategies developed during the last decade, and the cumulative 5-year overall survival rate of HCC remains unfavorable, probably due to its invasive behavior, as well as its histopathological and molecular heterogeneity that challenge molecular characterization and targeted therapeutic approaches. Moreover, the majority of patients are diagnosed in a more advanced stage, resulted in a much poorer prognosis. Thus, considerable work is still required to achieve a better understanding of the underlying mechanism at molecular level on the pathogenesis and carcinogenesis of HCC, which may be imperative for the development of robust biomarkers for early diagnosis and drug discovery.

In recent years, the rapid development of bioinformatics and emerging high-throughput techniques, such as microarray and next-generation sequencing (NGS), have enabled us to gain a comprehensive understanding of carcinogenesis and progression of various types of cancer. High-throughput platforms have been widely used in early diagnosis, histological identification, molecular classification, prognosis prediction, and drug resistance analysis of cancer [5–9]. Differentially expressed genes (DEGs), microRNAs (miRNAs), long noncoding RNAs (lncRNAs), and circular RNAs (circRNAs), as well as differentially methylated CpG sites have the potential to provide valuable clues to locate biomarkers in HCC. Nevertheless, data acquired from multiple studies may lead to false-positive results, considering the sample heterogeneity, different screening methods, different data mining approaches, and coupled effect of limited sample size in a single independent study. Integrated analysis based on collective datasets is a promising strategy to overcome these shortcomings. The microarray and sequencing data that deposited in the public databases such as The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), and International Cancer Genome Consortium (ICGC) have generated valuable information to identify biomarkers or to explore molecular landscapes of cancers, especially at transcriptome level [10–15]. Integrated transcriptome analyses have been employed to disclose the molecular mechanisms of cancers. However, few reports have been submitted to examine dysregulated genes and candidate biomarkers regarding HBV-related HCC (HBV-HCC) with combined datasets [16].

In the current study, seven datasets of gene expression profiles from GEO, TCGA, and ICGC were used to identify the hub genes that may play pivotal roles in the pathogenesis of HBV-HCC, comprising a total of 514 HBV-HCC tumor samples and 452 normal tissues. The common differentially expressed genes between tumors and adjacent normal liver tissues of four datasets (GSE84402 [17], GSE25097 [18], GSE94660 [19], and GSE121248 [20]) were screened by the R language software. The R package clusterProfiler [21] was utilized to conduct Gene ontology (GO) enrichment analysis and the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis to reveal the biological functions of DEGs. The protein-protein interaction (PPI) network of all common DEGs was constructed with information from the Search Tool for the Retrieval of Interacting Genes (STRING) database. Based on the entire network, hub gene identification and submodule analysis were performed using the Cyto-

Hubba plugin or the Molecular Complex Detection (MCODE) plugin of Cytoscape [22–24], respectively. The expression levels and correlation analysis of hub genes were then validated with the aid of complementary three datasets (TCGA, ICGC, and GSE14520). Furthermore, the diagnostic effectiveness of the hub genes was evaluated from TCGA and GSE14520, and the prognostic value was also determined in the GSE14520 dataset using univariate and multivariate analyses. Finally, Gene Set Enrichment Analysis (GSEA) and mRNA–microRNA interactions were operated to further clarify their possible molecular mechanisms. Taken all together, the crucial hub genes and associated key pathways identified from our findings would provide a potential indication for the molecular mechanism of HBV-HCC development and progression, and the particular proposed genes may shed additional insight into the early diagnosis, prognosis prediction, and therapeutic targets of HBV-HCC in the near future.

2. Materials and Methods

2.1. Gene Expression Dataset Acquisition. We chose four datasets of gene expression profiling from the GEO (<https://www.ncbi.nlm.nih.gov/geo/>) database to fetch the DEGs between HBV-HCC tumor samples and normal liver tissues, with the accession numbers of GSE84402, GSE25097, GSE94660, and GSE121248. GSE84402 and GSE121248 were both based on GPL570 (Affymetrix Human Genome U133 Plus 2.0 Array), while GSE25097 was based on GPL10687 (Merck Human RSTA Affymetrix 1.0 microarray, Custom CDF), and GSE94660 was based on GPL16791 (Illumina HiSeq 2500 (Homo sapiens)). HBV-HCC cases or samples were carefully filtered from all of the above studies. The dataset of GSE84402 contained 13 pairs of HBV-related hepatocellular carcinoma tissues and corresponding non-cancerous tissues. The dataset of GSE25097 comprised 73 HBV-HCC tissues and 67 paired adjacent nontumor samples. The dataset of GSE94660 consisted of 21 pairs of tumor and nonneoplastic liver tissues of HBV-HCC patients. The dataset of GSE121248 included 70 tissues from chronic hepatitis B-induced HCC and 37 adjacent normal tissues. In addition, we downloaded another three transcriptome datasets and corresponding clinical information for validation—GSE14520 from the GEO database including 213 HBV-HCC tissues and 220 normal samples, TCGA-LIHC data from The Cancer Genome Atlas (TCGA, <http://www.tcg.org/>) database including 70 HBV-HCC tissues and 49 normal samples, and ICGC-LIRI-JP data from the International Cancer Genome Consortium (ICGC, <https://icgc.org/>) including 53 HBV-HCC tissues and 45 paired normal samples. Table 1 summarized the detailed information of the selected datasets in this study.

2.2. Data Preprocessing and DEG Identification. All data were analyzed by the R software (version 3.6.0, <https://http://www.r-project.org/>). For the microarray data, we downloaded the raw data (.CEL files) and used the robust multiarray average algorithm [25] to conduct the background correction and quantile normalization; probes were subsequently matched with the corresponding gene name according to

TABLE 1: Detailed information of selected datasets in this study.

Study	Platforms	Etiology	Tumor	Normal	Total	Technology
GSE84402	GPL570	HBV	13	13	26	Microarray
GSE25097	GPL10687	HBV	73	67	140	Microarray
GSE94660	Illumina Hiseq	HBV	21	21	42	RNA-seq
GSE121248	GPL570	HBV	70	37	107	Microarray
GSE14520	GPL3921	HBV	213	220	433	Microarray
TCGA	Illumina Hiseq	HBV	71	49	120	RNA-seq
ICGC	Illumina Hiseq	HBV	53	45	98	RNA-seq

the platform annotation packages. When multiple probes were matched one the same gene, the highest expression value of all probes was selected to represent the final expression level. For RNA-seq data, we collected the normalized gene expression matrixes from the GEO database or HCCD database [26] directly. Meanwhile, the limma package [27] was applied to screen DEGs between HBV-HCC samples and noncancerous samples, with P value < 0.05 and $|\log_2 FC| \geq 1$ were set as the statistical cutoff criteria. The overlapping DEGs were identified by Venn plot. Besides, to minimize the chance for making type I error, we also adopted the integrating strategy by removing batch effect via “sva” [28] package for the three microarray datasets.

2.3. Functional Enrichment Analysis. To explore a further understanding of the potential biological function, GO and KEGG pathway enrichment analyses were performed on the overlapping dysregulated genes using the package “clusterProfiler” [21], and $P < 0.05$ was determined as a cutoff for significance.

2.4. PPI Network Construction. STRING database (version 11.0; <http://string-db.org/>), a free online tool to evaluate PPI information, was used to assess and integrate the interactive relationships among these DEGs. A combined confidence score of ≥ 0.7 (high confidence) was set as the threshold value. The Cytoscape software [24] (version 3.2.1; <http://www.cytoscape.org>) was then used to construct and visualize the PPI network. Besides, we also used a plugin from Cytoscape, the MCODE [23] (MCODE; version 1.5.1), to carry out the submodule analysis. For the identification of hub genes, we used the CytoHubba app [22] in Cytoscape with a combined method. Generally, nodes with top 60 scores of all the 11 algorithms in CytoHubba were picked up; then, the intersection based on more than 8 algorithms was taken as candidate hub genes, and two packages including UpSetR [29] and Venn detail [30] were used for visualization. Finally, the potential function of submodule genes or the hub genes was verified and displayed by the clusterProfiler [21] and the Gplot R packages [31].

2.5. Validation of Hub Gene Expression. As mentioned above, three datasets including GSE14520, TCGA-LIHC, and ICGC-LIRI-JP were used to verify the differential expression level of hub genes. The violin plots and heatmaps were drawn to compare the expression pattern between HBV-HCC tissues and noncancerous tissues, with the Wilcoxon test to

measure the statistical significance. The same three datasets were also used to perform the coexpression analysis of the selected hub genes. The correlation of hub genes’ expression and the clinical stages was investigated with the GSE14520 dataset and exhibited with boxplots.

2.6. Assessment of the Diagnostic and Prognostic Values. The receiver operating characteristic (ROC) analysis was conducted by means of R package pROC [32] to determine the power of potential hub genes in the diagnose of early phase and whole phase of HBV-HCC carcinogenesis, using their expression values from GSE14520 and TCGA-LIHC. The appropriate expression levels of the hub genes were served as cutoff values. For prognostic assessment, because of the sample size ($n < 100$) limitation of other datasets, only GSE14520 containing 213 HBV-HCC patients with complete OS data and sufficient clinicopathologic information were analyzed. In order to increase the robustness of the selection, we applied the “multi-split” strategy with log-rank test for 100 randomizations (75% portion of all samples were sub-sampled at each time) to evaluate the correlation between the OS and each hub gene expression level (Supplementary Figure 1). Those genes repeatedly showed significance for more than 75 times were considered as prognostic key genes and were further used to construct a linear combination as the risk signature with the following formula: risk score = sum (coef (k) \times expression value of k), where k represents the candidate prognostic key genes. All patients were divided as high-risk or low-risk groups according to the median risk score. Kaplan-Meier method with a log-rank test was used to compare the survival curves by using the survival package [33]. The time-dependent ROC was depicted to estimate the predictive ability of the risk signature for patients’ OS survival. The associations of clinicopathologic features and the risk signature were determined by Pearson chi-square test or Fisher’s exact test. Univariable and multivariable Cox regression analyses were conducted to identify independent prognostic factors. For all statistical tests, $P < 0.05$ was set as significant cutoff.

2.7. Gene Set Enrichment Analysis. The genes that differentially expressed and associated with risk stratification of OS survival were analyzed by GSEA, which was employed to examine the statistical significance of a priori defined set of genes between different phenotypes. In the current study, we focused on the KEGG pathways (c2.cp.kegg.v7.1) and molecular function of GO gene sets (c5.mf.v7.1). 1000 times

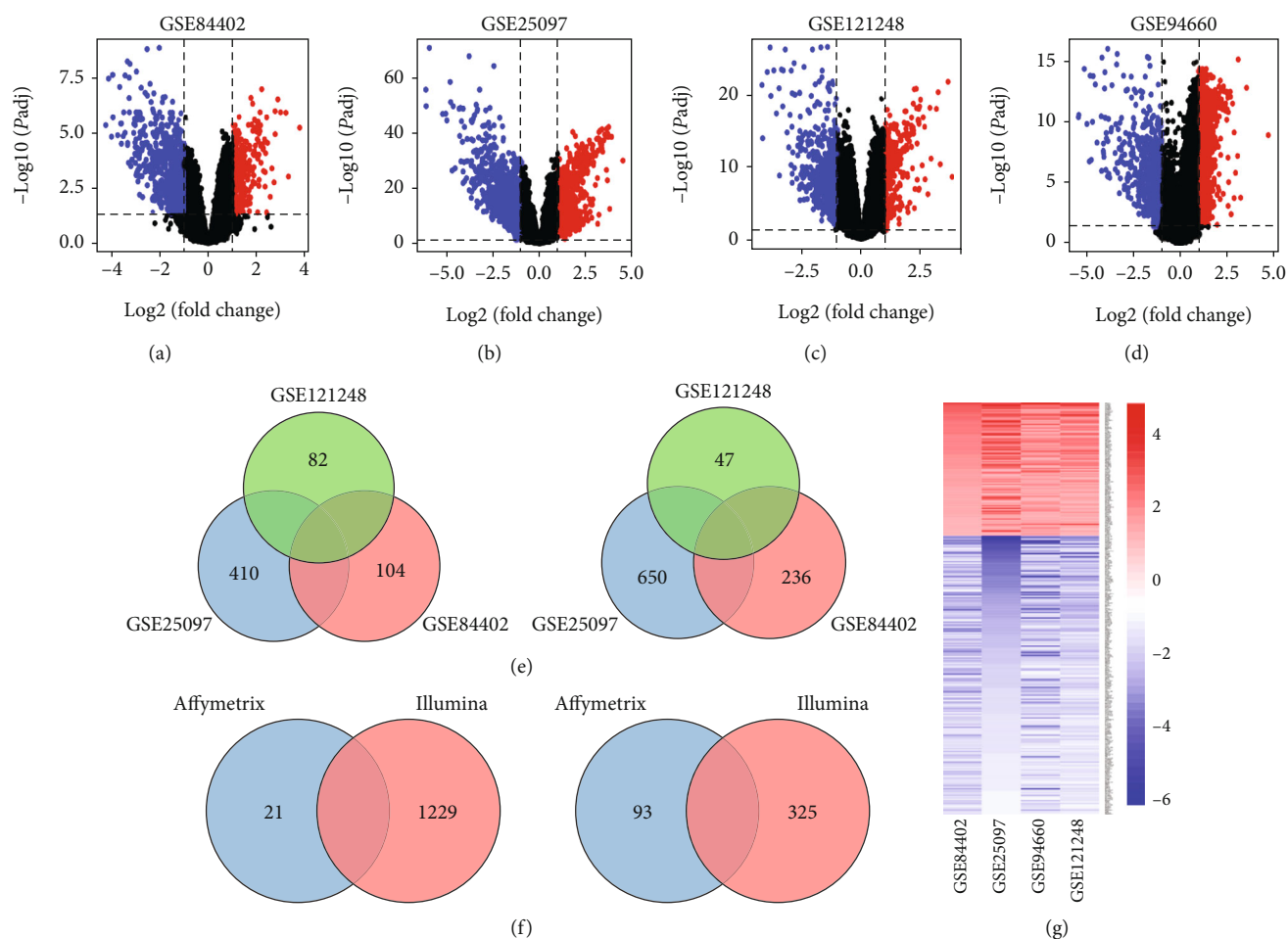


FIGURE 1: Identification of overlapping DEGs among the screening datasets of HBV-HCC. (a–d) Volcano plot for GSE84402, GSE25097, GSE121248, and GSE94660. (e) Venn diagram for the overlapping DEGs from Affymetrix biosystems. (f) Common DEGs shared by Affymetrix and Illumina platforms. (g) Heatmap of 309 common DEGs. Blue represents downregulated genes while red represents upregulated genes. Each column represents one dataset, and each row represents one gene. DEGs: differentially expressed genes; HBV-HCC: HBV-related hepatocellular cancer.

of permutations were performed to obtain the enrichment score (ES) and the normalization enrichment score (NES) for each gene set. Significant gene sets were identified with a nominal P value < 0.05 , combined with an FDR $< 25\%$ and normalized enrichment score (NES) > 1 or < -1 .

2.8. miRNA-mRNA Interaction Prediction. We predicted the liver-specific miRNA-mRNA interaction of all hub genes using an online tool—miRNet [34], which integrated several well-annotated databases, including miRTarBase v7.0, TarBase v7.0, and miRecords. miRNA-mRNA correlation network was constructed with the Cytoscape software. In the network, diamond nodes denoted the mRNAs while the rectangle nodes represented the miRNAs.

3. Results

3.1. Identification of DEGs in HBV-HCC. With the filtering criteria mentioned above, DEGs in HBV-HCC carcinogenesis were achieved using a total of 315 clinical samples from four GEO datasets (GSE25097, GSE84402, GSE121248, and

GSE94660). In all, there were 623 upregulated and 1144 downregulated genes identified from GSE25097, 319 upregulated and 775 downregulated genes from GSE84402, 319 upregulated and 564 downregulated genes from GSE121248, and 1329 upregulated and 534 downregulated genes from GSE94660. The volcano plots of all DEGs in each of the four datasets were shown in Figures 1(a)–1(d). For overlapping analysis of the Venn diagram, 121 upregulated genes or 302 downregulated genes were firstly shared by Affymetrix biosystems (Figure 1(e)); then, 309 common DEGs were obtained by Affymetrix and Illumina platforms, consisting of 100 upregulated genes and 209 downregulated genes (Figures 1(f) and 1(g)). To increase the robustness of these 309 DEGs, we conducted the integrating analysis of the three microarray datasets. After removing the batch effect, all 309 DEGs were still showed significant (Supplementary Figure 2, Supplementary Table 1).

3.2. DEG Function Analysis. Three categories of GO comprising of biological process (BP), molecular function (MF), and cellular component (CC), together with KEGG enrichment,

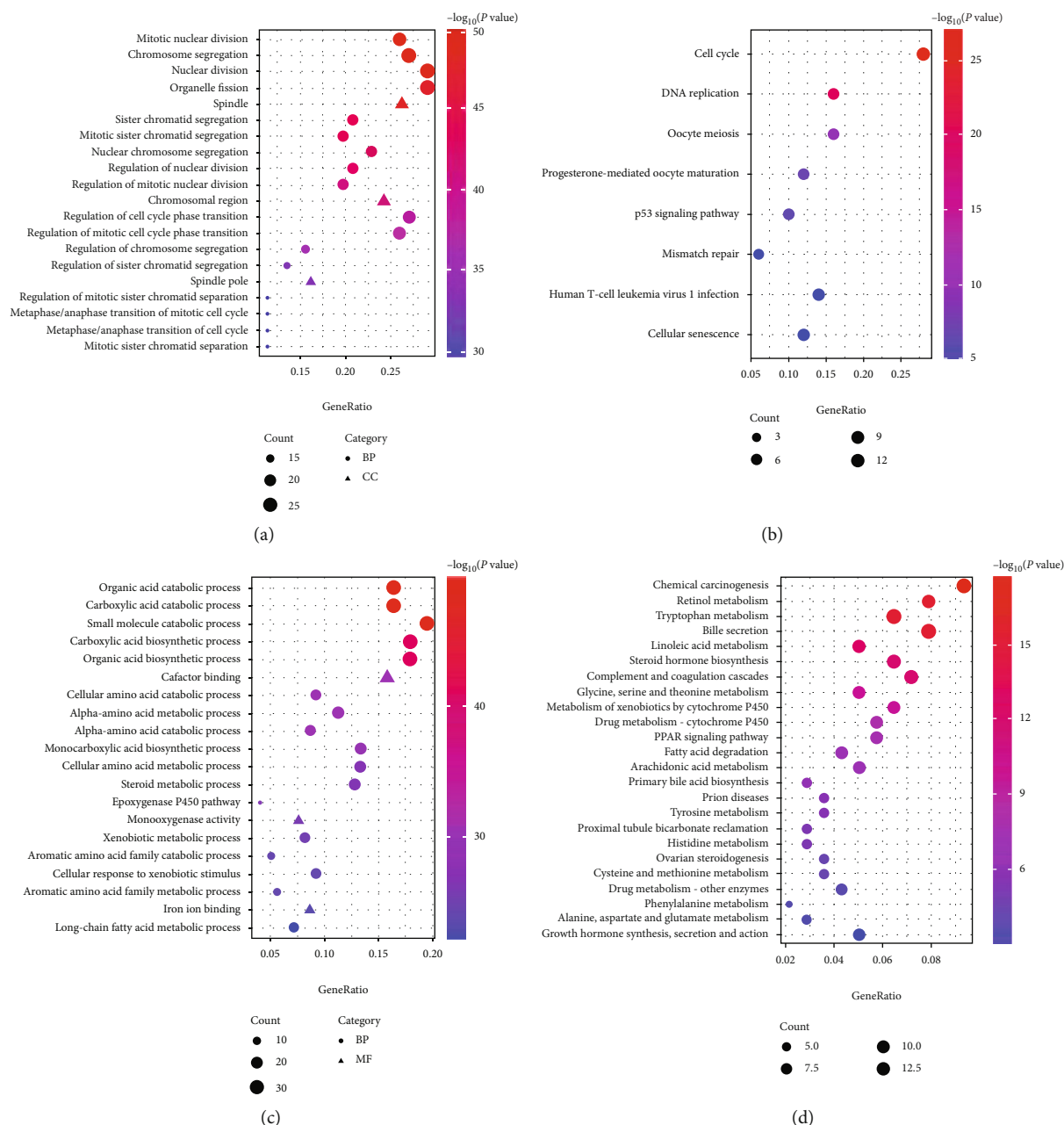


FIGURE 2: Function enrichment analysis of the overlapping DEGs. (a) GO analysis for the upregulated DEGs (top 20 GO terms are shown). (b) KEGG pathway enrichment for the upregulated DEGs. (c) GO analysis for the downregulated DEGs (top 20 GO terms are shown). (d) KEGG pathway enrichment for the downregulated DEGs. BP: biological process; CC: cellular component; MF: molecular function.

were performed for the upregulated and downregulated genes. For GO functions analysis, upregulated genes were involved in multiple GO terms, such as mitotic nuclear division, chromosome segregation, nuclear division, and spindle. On the other hand, significant GO terms that associated with downregulated genes were organic acid catabolic process, carboxylic acid catabolic process, small molecule catabolic process, carboxylic acid biosynthetic process, and so on. For KEGG pathway analysis, the upregulated genes were mostly enriched in the cell cycle, DNA replication, oocyte meiosis, progesterone-mediated oocyte maturation, p53 signaling pathway, and mismatch repair. Meanwhile, the downregu-

lated genes were mainly related to chemical carcinogenesis, retinol metabolism, tryptophan metabolism, bile secretion, and linoleic acid metabolism (Figure 2).

3.3. PPI Network Construction. For the DEG interaction inspection, we constructed a PPI network of 209 nodes and 1676 edges by the STRING database and Cytoscape software, with a strict criterion (combined interaction score ≥ 0.7). The PPI network contained 84 upregulated DEGs and 125 downregulated DEGs (Supplementary Figure 3). The clustering coefficient was 0.525, and the average degree was 16.038. In addition, the MCODE app of Cytoscape detected the most

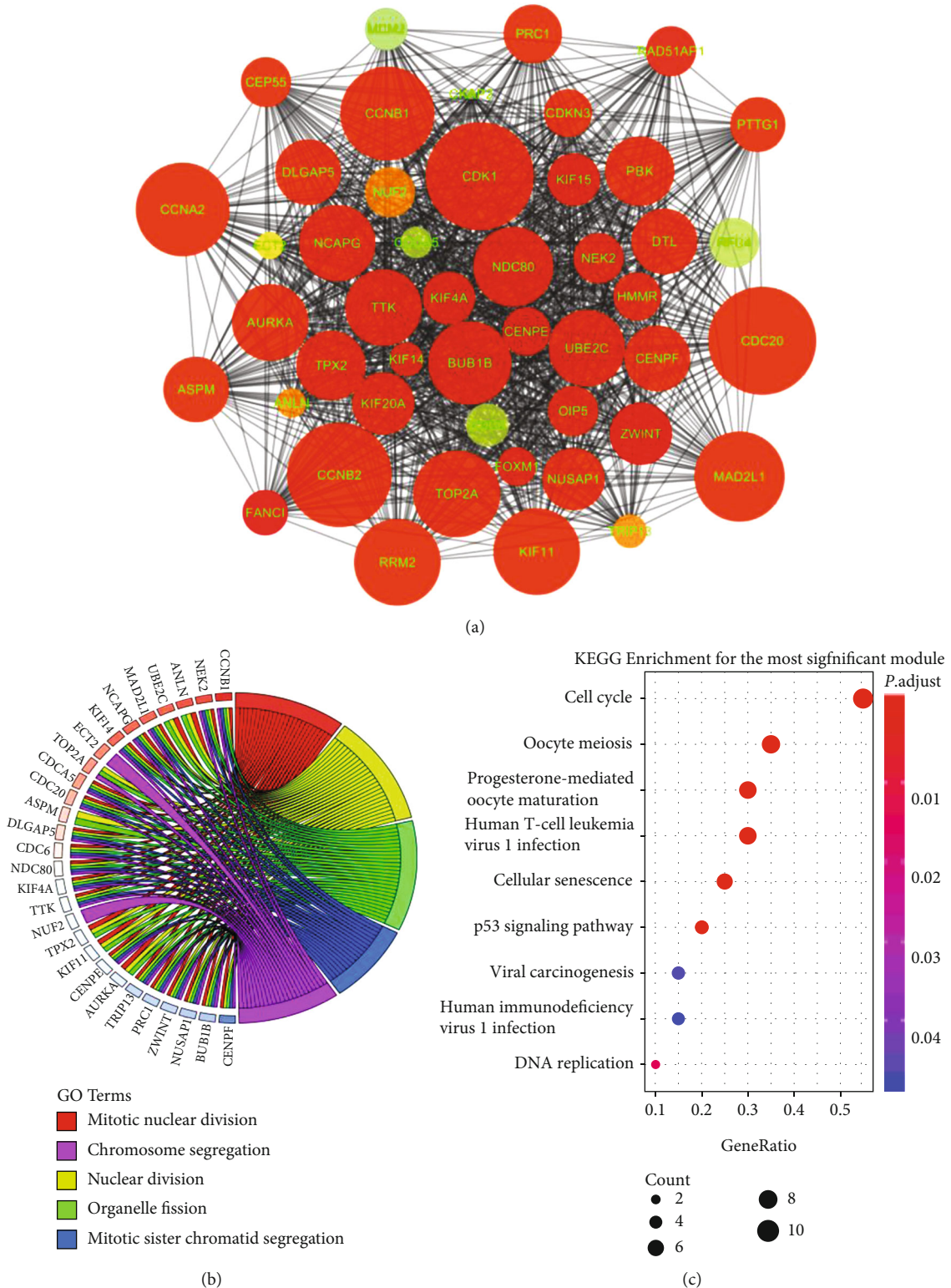
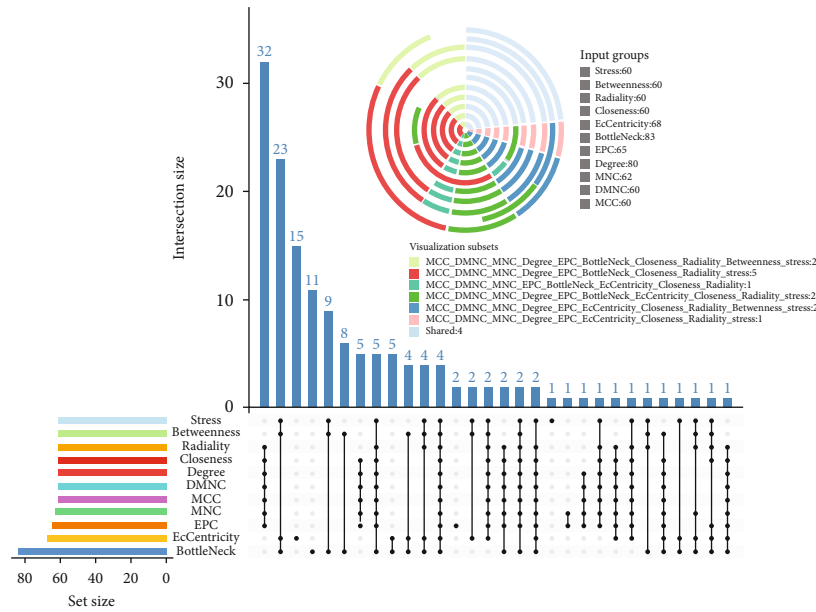


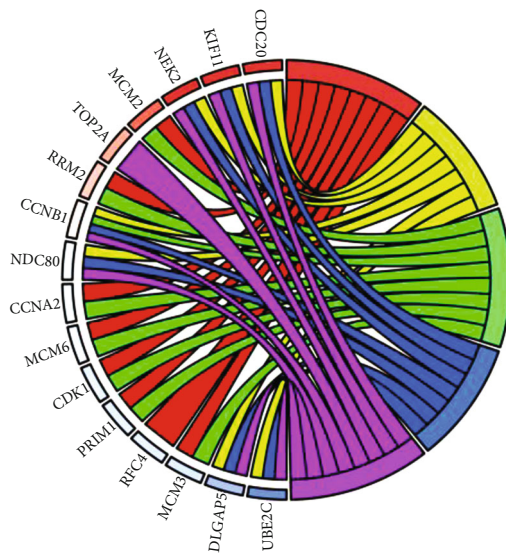
FIGURE 3: Subnetwork analysis of the DEGs PPI. (a) The most significant module selected by MCODE plugin (MCODE score > 40), comprising 47 nodes. (b) Top five related GO terms enriched by the DEGs in the module. (c) KEGG enrichment analyses for the DEGs in the module.

significant module with a high network score (>40), which consisted of 47 nodes and 1017 edges (Figure 3(a)). Functional analysis revealed that mitotic nuclear division and cell cycle were the most significant GO and KEGG pathway enriched by the module (Figures 3(b) and 3(c)).

For the hub gene identification, we used a combined method by CytoHubba. Interestingly, all of the 17 hub genes were upregulated DEGs, and most of them (UBE2C, RRM2, RFC4, TOP2A, MCM2, CDK1, CCNB1, HMMR, CDC20, CCNA2, NEK2, NDC80, DLGAP5, and KIF11) were

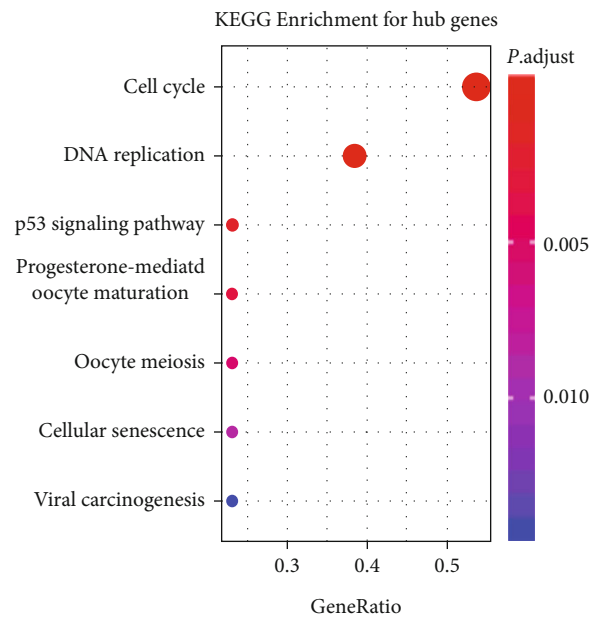


(a)



- GO Terms
- DNA replication
 - Nuclear division
 - Cell cycle G1/S phase transition
 - Regulation of mitotic nuclear division
 - Regulation of nuclear division

(b)



(c)

FIGURE 4: Hub gene identification and functional analysis. (a) The combination of Upset plot and Vennpie plot shows the 17 hub genes identified by CytoHubba plugin through DEG PPI network, with an overlapping strategy. (b) Top five related GO terms of the hub genes. (c) Result of KEGG pathway analysis of the hub genes.

involved in the significant module, while only 3 hub genes (MCM6, MCM3, and PRIM1) were not included (Figure 4(a)). We also employed the GO and KEGG analyses of all the hub genes for further validation of their biological functions, such as DNA replication and cell cycle regulators (Figures 4(b) and 4(c)).

3.4. *Hub Gene Validation.* To validate the differentially expressed levels of the selected hub genes, three other datasets were used in this study. Consistent with the screening cohorts from the GEO database, all of the 17 definitive hub genes showed the significantly higher expression trends between HBV-HCC cancer samples and adjacent liver

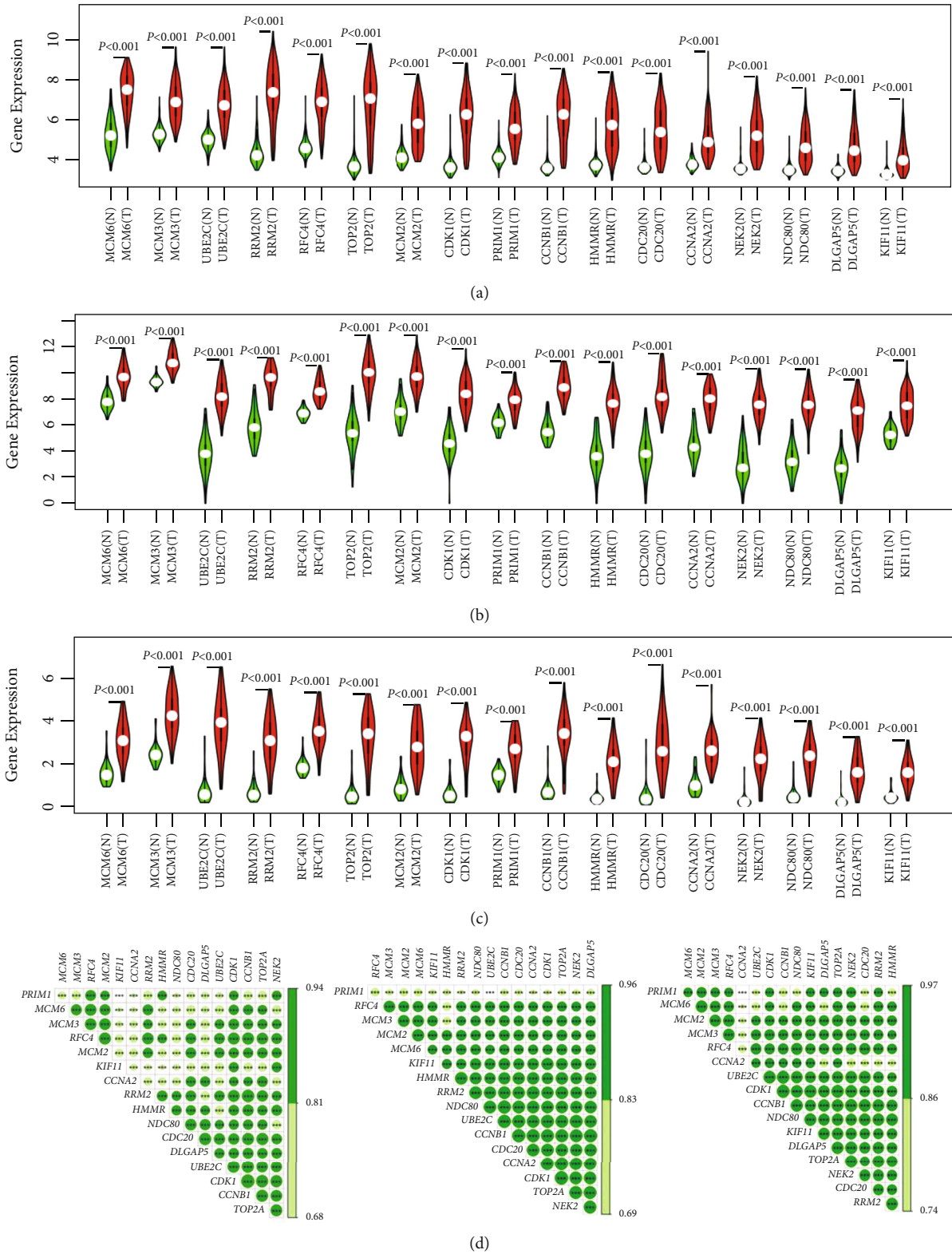


FIGURE 5: Validation of the aberrant expression levels for the selected hub genes and coexpression analysis. (a–c) Violin plots showing the significantly increased expression values for all of the 17 hub genes based on GSE14520, TCGA-LIHC, and ICGC-LIRI-JP. (d) Pearson correlation analysis among expression levels of the 17 hub genes for GSE14520, TCGA-LIHC, and ICGC-LIRI-JP. The color depth indicates the degree of correlation. The darker the color, the higher the correlation coefficient. *** $P < 0.001$.

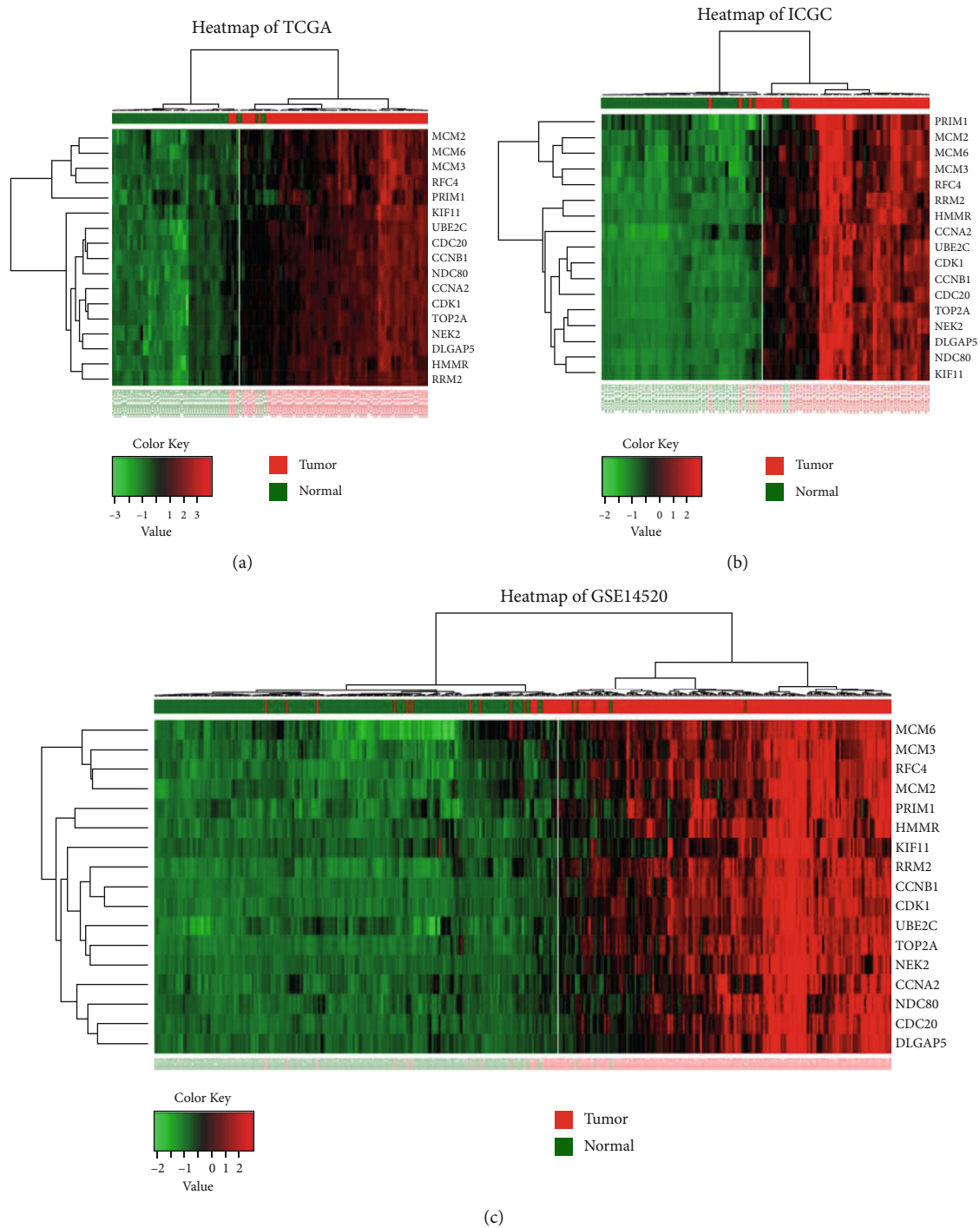


FIGURE 6: Clustering heatmaps of 17 hub genes based on (a) TCGA-LIHC, (b) ICGC-LIRI-JP, and (c) GSE14520. Red denotes high expression levels while green denotes low expression levels.

normal ones (Figures 5(a)–5(c)). Moreover, the hierarchical clustering of the hub genes and samples based on the three datasets revealed their good potential in discriminating the tumors from healthy tissues (Figure 6). Heatmaps of Pearson correlation suggested the high correlations of these hub genes in all three datasets, supporting the underlying hypothesis that hub genes may strongly interact with each other and play critical roles in the development of HBV-HCC (Figure 5(d)).

To seek for the clinical relevance of these hub genes, we analyzed the transcription expression levels of the hub genes according to patients' TNM stages and their BCLC stages. Consequently, most of the hub genes showed no statistical difference between TNM stage I and other stages, which suggested their predominant roles in the initiation of carcinogenesis, but they may not good indicators for HBV-HCC progression (Figure 7). A similar result was reached based

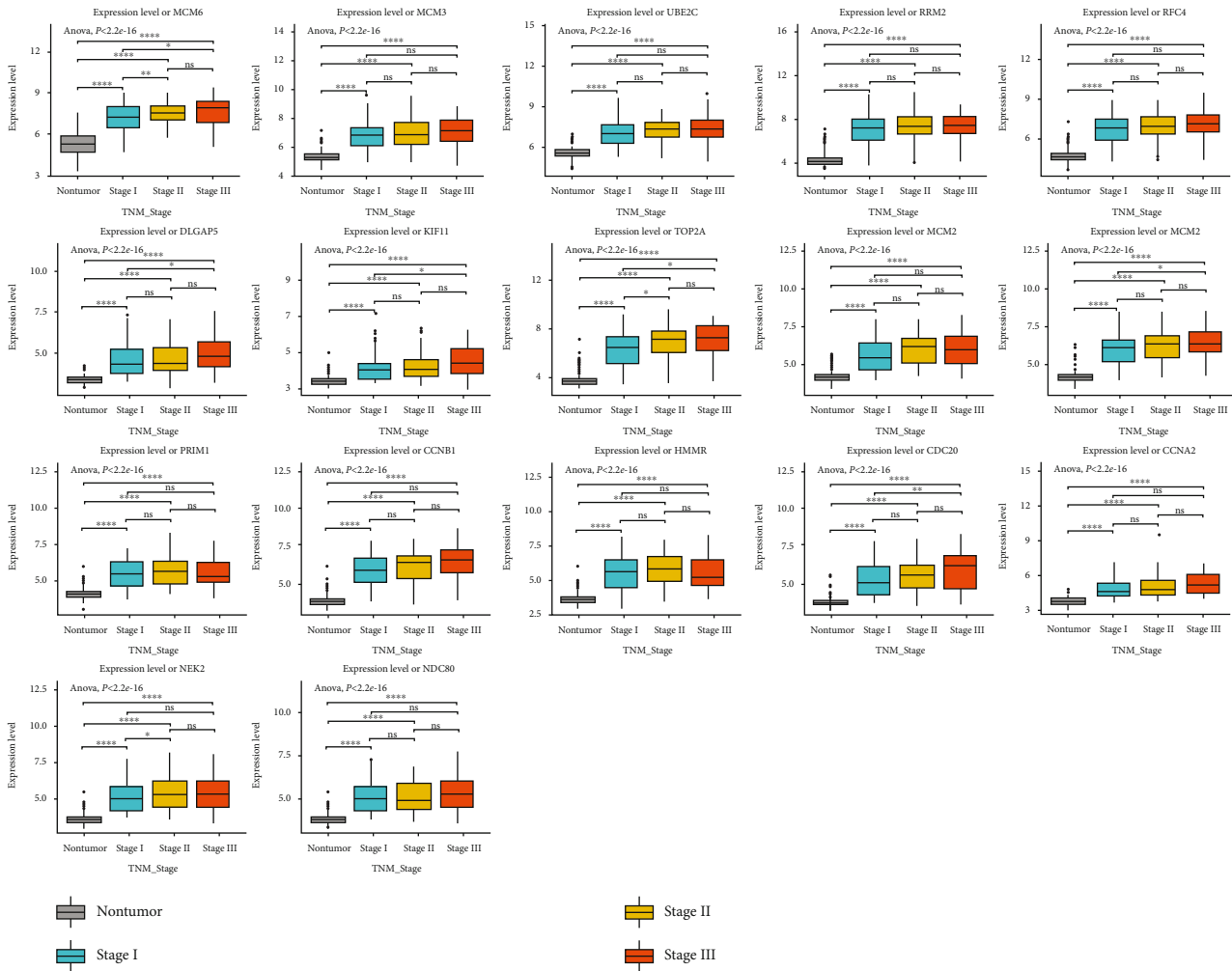


FIGURE 7: Boxplots showing the relative expression levels of 17 hub genes across normal liver tissues and cancer tissues with different TNM stages for HBV-HCC. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, and **** $P < 0.0001$.

on the BCLC staging system, which implied that only MCM6, RRM2, and CCNB1 might be associated with HBV-HCV progression (Supplementary Figure 4).

3.5. Diagnostic Value Assessment. The above findings prompted us to speculate that these hub genes may have good diagnostic efficiency for HBV-HCC, which was verified by plotting ROC curves using TCGA-LIHC and GSE14520. As the result, the area under the curves (AUCs) of all hub genes ranged from 0.91 to 0.99 (Supplementary Figure 5), indicating their excellent diagnostic values for distinguishing tumor tissues and adjacent normal ones. Next, we were curious whether these hub genes also play a role in the early detection of HBV-HCC, which was even more crucial for clinical intervention. Thus, we especially focused on early stage (TNM stage I) cases, and ROC analysis of individual hub genes proved their great potential in the early diagnosis of HBV-HCC (Figure 8).

3.6. Survival Analysis. To elucidate the prognostic values of the hub genes, we performed the OS survival analysis with the cohort from GSE14520, including 213 HBV-HCC

patients. The resample-based log-rank test resulted in two robust prognostic hub genes (with high repentance frequency of >0.75 showing significant during resampling): TOP2A and KIF11. Then, a prognostic signature was built with these two hub genes, and all patients were assigned to high- or low-risk group based on the median risk score (0.997189). The result of risk score for each patient was shown in Figure 9(a). Figure 9(b) suggested the significant difference of risk scores between high- or low-risk group, and OS survival rate was significant higher in low-risk group by chi-square analysis (Figure 9(c)). As for prediction accuracy, we plotted the time-dependent roc curves by risk score, and the risk signature showed the AUC values at 1, 3, and 5 years was 0.626, 0.643, and 0.693, respectively (Figure 9(d)). For Kaplan-Meier survival analyses, both the OS survival rate and the recurrence-free survival rate were shown to be significantly higher in low-risk group than high-risk group (Figures 9(e) and 9(f)).

Furthermore, we stratified patients into different risk subgroups by several clinicopathologic parameters (age, gender, ALT, AFP, level, main tumor size, cirrhosis, BCLC stage, TNM stage, and CLIP stage) using the two-hub gene-

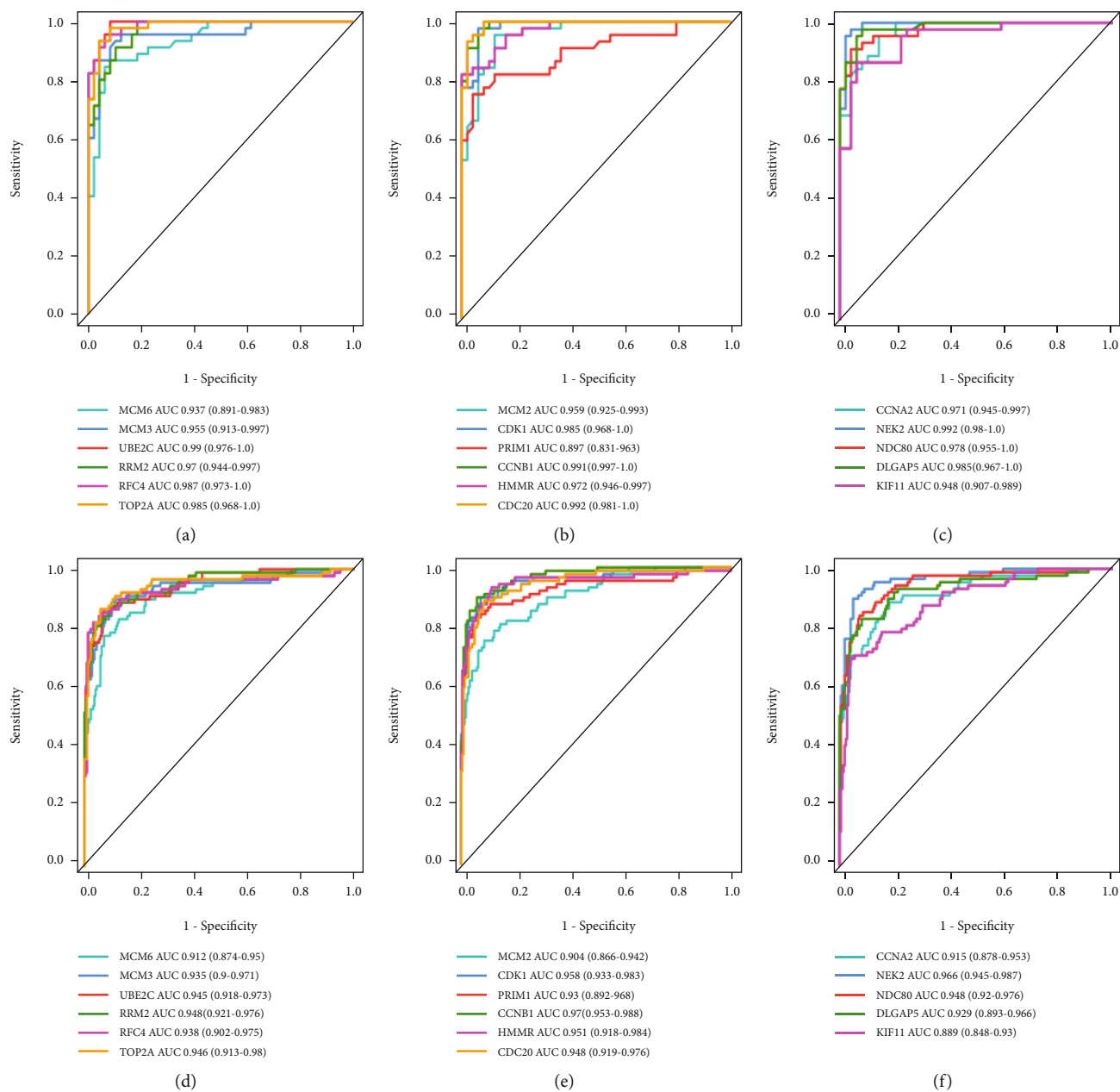


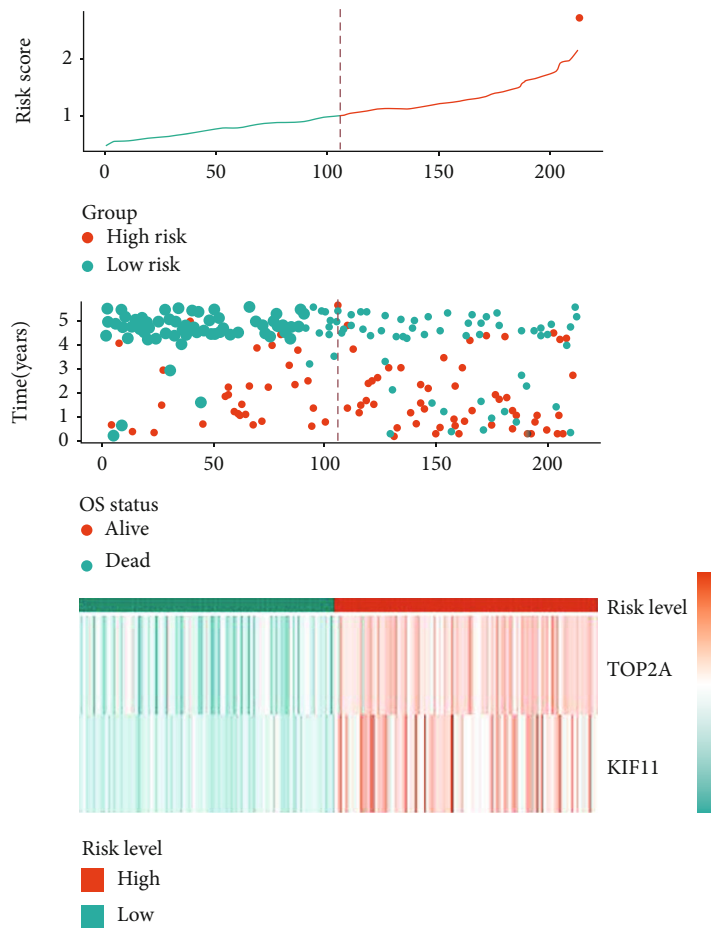
FIGURE 8: The ROC curves and AUC (95% CI) for each of the selected hub genes to evaluate their efficiency in the early diagnosis of HBV-HCC based on (a-c) TCGA-LIHC cohort and (d-f) GSE14520 cohort. Colored lines denote sensitive curves for each hub gene, and grey line denotes the identify line. ROC: receiver operating characteristic; AUC: area under the curve.

based signature. Interestingly, results showed that our classifier was still statistically significant in most subgroups (Supplementary Figure 6), suggesting its good potential and possible application to add prognostic value to the existing staging systems.

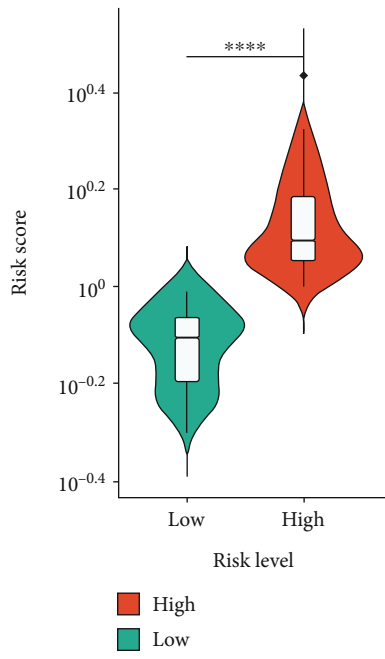
Moreover, risk levels based on the two-hub gene-based signature also suggested to be significantly associated with other aggressive clinicopathological parameters, such as TNM stage ($P = 0.002$), BCLC stage ($P = 0.013$), CLIP stage ($P = 0.003$), and alpha fetal protein (AFP) level ($P = 0.003$) by chi-square test (Table 2). Multivariate Cox regression analysis indicated that after adjusting for main tumor size,

cirrhosis, TNM stage, BCLC stage, CLIP stage, and AFP level, the risk signature was still significantly correlated with OS survival outcome (HR = 1.807, 95%CI = 1.126 - 2.899, and $P = 0.014$), implying that the two-hub gene signature served as an independent prognostic factor for HBV-HCC patients (Table 3).

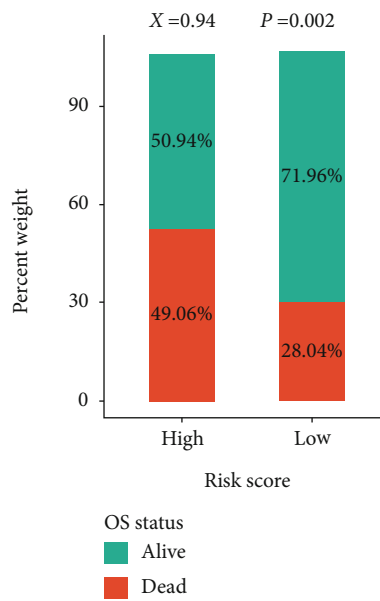
3.7. Gene Set Enrichment Analysis. GSEA is a powerful statistical approach to identify classes of genes that are significantly associated with different disease phenotypes. Thus, inspired by the results of the prognostic analysis, we operated the GSEA to investigate the molecular mechanisms between



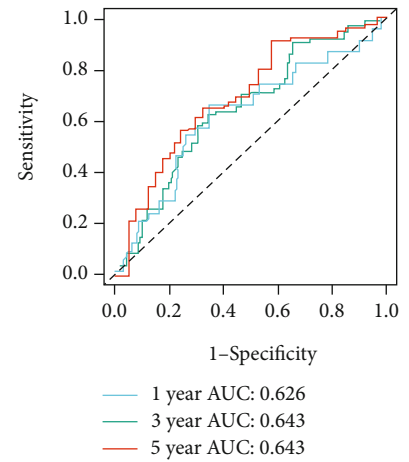
(a)



(b)



(c)



(d)

FIGURE 9: Continued.

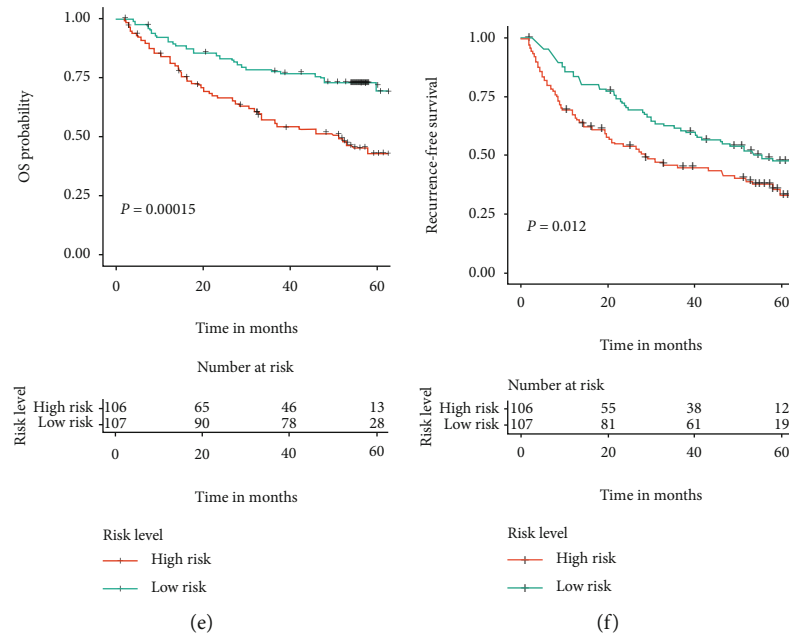


FIGURE 9: Prognostic assessments of hub genes by GSE14520 cohort. (a) Risk score, survival outcome, and hub gene expression values for each patient in high- and low-risk groups. (b) Comparison of risk scores between high- and low-risk groups. **** $P < 0.0001$. (c) Distribute of different survival status in high- and low-risk score groups. Statistical significance was determined by chi-square test. (d) The time-dependent ROC curves for the two-hub gene-based signature at 1, 3, and 5 years to assess accuracy of prognostic prediction. (e, f) Kaplan-Meier curves of (e) OS and (f) RFS in HBV-HCC patients based on the risk score classification. P was calculated by the log-rank test, and $P < 0.05$ was considered statistically significant. HBV-HCC: HBV-related HCC; OS: overall survival; RFS: relapse-free survival.

high- and low-risk groups divided by the risk signature. GSEA was performed by KEGG at first, and top significant pathways were identified as KEGG_spliceosome ($P < 0.001$ and NES = 2.149), KEGG_cell_cycle ($P < 0.001$ and NES = 1.959), and KEGG_oocyte_meiosis ($P < 0.001$ and NES = 1.934) (Figure 10(a)). For the gene set distribution from molecular function component of the GO database, GO_translational_initiation ($P < 0.001$ and NES = 2.233), GO_meiotic_cell_cycle ($P < 0.001$ and NES = 2.191), and GO_RNA_splicing_via_transesterification_reactions ($P < 0.001$ and NES = 2.149) were ranked as the most significant terms (Figure 10(b), Supplementary Tables 2–3). These results suggested that this risk signature may exert a poorer survival for HBV-HCC patient via known crucial cancer pathways.

3.8. Prediction of miRNA-mRNA Interaction Network. miRNAs have been extensively documented to regulate tumorigenesis at transcriptome level or posttranscriptional level in various cancers. Thus, we predicted the candidate miRNAs that may target these hub genes in human liver by using the miRNet online platform. Then, the miRNA-hub gene interaction network was established by Cytoscape (Supplementary Figure 7). There were 55 nodes and 77 edges involved in the network, including 16 hub genes (NDC80 was excluded) and 39 miRNAs. In the network, PRIM2, CCNA2, and RRM2 were recognized as the top 3 hub genes that had most neighbors of miRNAs, while hsa-mir-34a-5p, hsa-mir-192-5p, and hsa-mir-24-3p were the top 3 miRNAs with most targeted hub genes. The miRNA-hub gene network based on their regulatory relationships may

provide a forceful basis for the further exploration of the molecular mechanisms of HBV-HCC.

4. Discussion

Despite the great advances in clinical management and remarkable progress in understanding the pathogenesis of HCC, the incidence and mortality rates of this malignant cancer remain unacceptably high. Chronic hepatitis B is the primary etiological factor for HCC in China and other parts of Asia [4]. With the identification of diagnostic and prognostic biomarkers of HBV-HCC, we have attempted to provide valuable insight into the molecular mechanism of HBV-HCC during tumorigenesis and development.

Public databases like GEO, TCGA, and ICGC that deposit massive datasets of high-throughput technologies like microarray and NGS platforms have facilitated the strategies for mining of integrated data, which could overcome the limitations of the small sample size in one individual cohort and heterogeneity among different studies. In the present study, we conducted DEG screening based on the transcription profiling data of GSE84402, GSE25097, GSE94660, and GSE121248 at first, and there were 309 overlapped DEGs were identified between the tumor and normal patients, comprising 100 upregulated genes and 209 downregulated genes. The result of GO analysis revealed that the upregulated DEGs were significantly enriched in the regulation of cell division activities (such as mitotic nuclear division, chromosome segregation, and nuclear division), while the downregulated DEGs were closely related to

TABLE 2: Clinicopathological features of HBV-HCC patients according to the two-hub gene-based signature in the cohort of GSE14520.

Characteristics	Low risk (<i>n</i> = 107)	High risk (<i>n</i> = 106)	<i>P</i> value
Gender			
Female	17	13	0.447
Male	90	93	
Age			
<50	48	47	0.939
≥50	59	59	
ALT level			
High (>50 U/L)	41	47	0.518
Low (≤50 U/L)	66	59	
Main tumor size [†]			
Large (>5 cm)	35	40	0.412
Small (≤5 cm)	72	65	
Cirrhosis			
Yes	96	100	0.213
No	11	6	
TNM stage			
I	56	33	0.002*
II+III	51	73	
BCLC stage			
0-A	90	74	0.013*
B-C	17	32	
CLIP stage			
0	58	36	0.003*
≥1	49	70	
AFP level [†]			
High (≥300 ng/ml)	37	58	0.007*
Low (<300 ng/ml)	67	49	

[†]The information was missing for certain patients. **P* < 0.05 by χ^2 test. HBV-HCC: HBV-related HCC; ALT: alanine aminotransferase; AFP: alpha fetoprotein; TNM: tumor-node-metastasis; BCLC: Barcelona Clinic Liver Cancer; CLIP: Cancer of the Liver Italian Program.

multiple cellular “catabolic process” and “biosynthetic processes” (such as organic acid catabolic process and carboxylic acid biosynthetic process). KEGG pathway analysis revealed that the upregulated DEGs were mainly involved in the cell cycle, DNA replication, oocyte meiosis, and others. Meanwhile, the downregulated DEGs were relevant to chemical carcinogenesis, retinol metabolism, tryptophan metabolism, bile secretion, and so forth.

After that, we utilized a combined strategy to identify the 17 hub genes (MCM6, MCM3, UBE2C, RRM2, RFC4, TOP2A, MCM2, CDK1, PRIM1, CCNB1, HMMR, CDC20, CCNA2, NEK2, NDC80, DLGAP5, and KIF11) from the PPI network established by the STRING database. Most of these hub genes were previously reported as oncogenes, therapeutic targets, or potential biomarkers in HCC [35–51]. We then validate the dysregulated mRNA expression levels of these hub genes using GSE14520, TCGA-LIHC, and ICGC-

LIRI-JP datasets. The positive coexpression relationships of the hub genes were proved by Pearson’s correlation analysis, implying the highly active interactions during the tumorigenesis. Boxplots were used to demonstrate the relevance between gene expression levels and pathological stages; however, most of the hub genes showed no significant association between early stages and late stages, stimulating us to propose that they may be used in early diagnosis for HBV-HCC. As the consequence, the ROC curves of all the 17 hub genes showed high diagnostic values for TNM stage I patients and adjacent normal tissues, suggesting their good potential in further exploiting early diagnosis, including related miRNAs, circRNAs, and aberrantly methylation markers that based on these hub genes.

For the overall survival analysis, with the aid of cohort from GSE14520, we established a risk signature with two prognostic hub genes: TOP2A and KIF11, which was demonstrated to be an independent prognostic predictor for HBV-HCC patient by univariate and multivariate analyses and was significantly correlated with tumor staging systems and AFP levels. Our GSEA result also revealed key molecular functions and KEGG pathways (especially for gene sets related to cell cycle) that involved in carcinogenesis that may be associated with OS survival stratification by the risk signature for HBV-HCC patients.

TOP2A, encoded by TOP2A gene, is a DNA topoisomerase that participates in many processes during transcription and replication through altering DNA topological structure. Previous studies confirmed that the aberrant TOP2A expression was observed in various cancer subtypes, such as breast [52], colon [53], ovarian [54], gastric [55], prostate cancer [56], and HCC [40]. In the current study, TOP2A showed high expression value in HBV-HCC, which agreed with previous results. Recent studies substantiated its oncogenic role during the tumorigenesis and development of many malignancies [52, 53]. For example, Zhang et al. found that knock-down of TOP2A could induce apoptosis and suppress cell proliferation and invasion via Akt and ERK signaling pathways in colon cancer [53]. These findings strongly imply that TOP2A may be served as an anticancer therapeutic target for clinical treatment. Actually, several TOP2A inhibitors have been approved by the US Food and Drug Administration [57], and other compounds were tested in multiple trials [58, 59]. A study on adrenocortical carcinoma (ACC) manifested that aclarubicin was the best agent of 14 TOP2A inhibitors that can decrease proliferation and tumor spheroid size in locally advanced and metastatic ACC [58]. But candidate TOP2A inhibitors with high efficacy for HCC were still rare. Considerable effort was required to explore effective reagents for HCC. Furthermore, early discoveries revealed that the elevated TOP2A expression implicated the worse overall survival for multiple cancers [52, 54, 56]. In accordance with these findings, our present study demonstrated that upregulation of TOP2A was closely related to the poor outcome for HBV-HCC patients.

KIF11 or as BimC, Eg5, belonging to kinesin superfamily, which function as nanomotors to mediate various kinds of spindle dynamics, is well known to play an essential role during cell mitosis, including chromosome positioning, bipolar

TABLE 3: Univariate and multivariate analyses of the cohort GSE14520.

Variables	Group	Univariate analysis			Multivariate analysis		
		HR	95% CI	P value	HR	95% CI	P value
Risk level	High/low	2.293	1.461-3.600	<0.001 [†]	1.807	1.126-2.899	0.014*
Gender	Male/female	0.599	0.289-1.242	0.168	-	-	-
Age	≥50/<50	1.190	0.771-1.837	0.432	-	-	-
ALT level (U/L)	≥50/<50	1.107	0.716-1.712	0.648	-	-	-
Main tumor size (cm)	≥5/<5	2.063	1.330-3.199	0.001 [†]	1.269	0.77-2.091	0.349
Cirrhosis	Yes/no	0.224	0.055-0.910	0.036 [†]	0.349	0.084-1.440	0.145
TNM stage	II+III/I	0.348	0.210-0.577	<0.001 [†]	0.579	0.326-1.027	0.061
BCLC stage	B-C/0-A	0.283	0.18-0.445	<0.001 [†]	0.478	0.27-0.848	0.012*
CLIP stage	≥1/0	2.133	1.338-3.398	0.001 [†]	1.129	0.549-2.322	0.741
AFP level (ng/ml)	≥300/<300	1.572	1.019-2.426	0.041 [†]	1.006	0.541-1.869	0.986

[†]Significant in univariate Cox regression and were further enrolled for multivariable Cox regression analysis. * $P < 0.05$ by both univariate and multivariate analyses. AFP: alpha fetoprotein; TNM: tumor-node-metastasis; ALT: alanine aminotransferase; BCLC: Barcelona Clinic Liver Cancer; CLIP: Cancer of the Liver Italian Program.

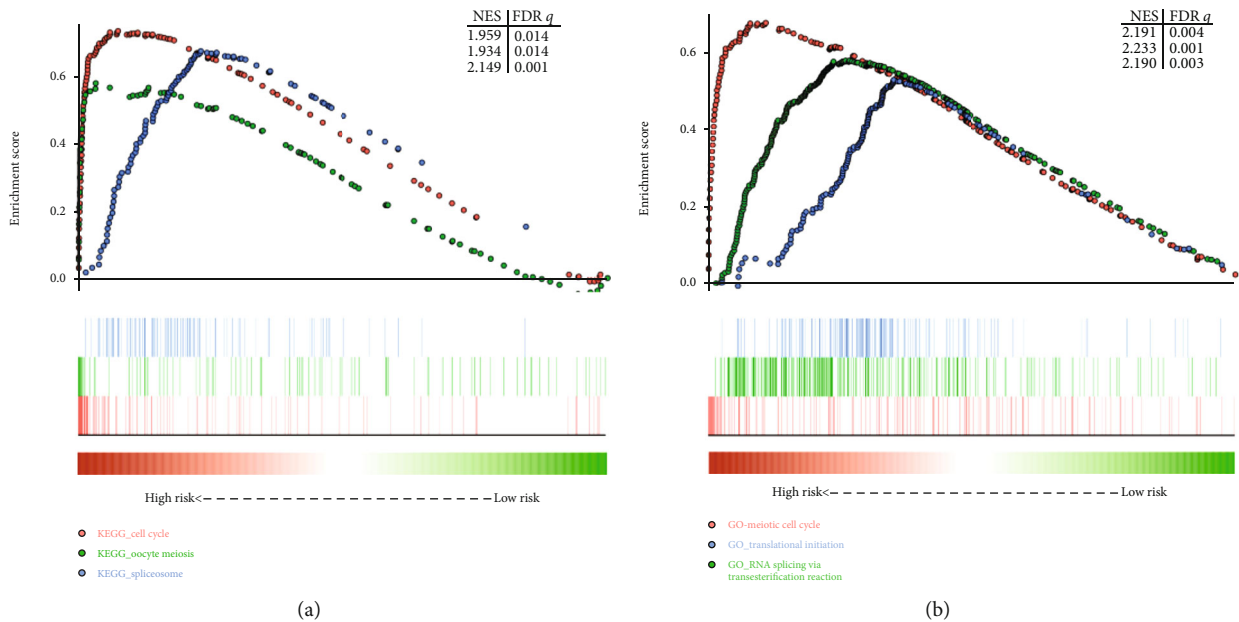


FIGURE 10: GSEA results of high- or low-risk groups divided by the two-hub gene-based signature in GSE14520 cohort. (a) Results of GO terms enriched in highly risk group vs. low-risk group. (b) Results of KEGG pathways enriched in highly risk group vs. low-risk group. GSEA: Gene Set Enrichment Analyses.

spindle formation and maintenance, and antiparallel microtubule sliding, as well as microtubule crosslinking [60–62]. It may also increase translational efficiency by mediating the association of ribosomes and microtubules [63]. Recently, KIF11 has been reported to be a novel potential candidate prognostic biomarker or therapeutic target in human cancers including breast cancer [64], ovarian cancer [65], oral cancer [66], peripheral nerve sheath tumors [67, 68], and lung cancer [69]. In line with previous findings, we identified KIF11 as an oncogene during tumorigenesis of HBV-HCC, and its expression level was significantly higher in tumor samples compared with adjacent normal tissues. It was also proved to hold good potential for early detection of HBV-HCC. For prognostic analysis, KIF11 and TOP2A performed

jointly well in predicting prognosis by multivariate regression. In fact, a growing body of well-known KIF11 inhibitors such as monastrol [70], S-trityl-L-cysteine (STLC) [71], HR22C16 [72], and CK0106023 [73] have been extensively studied, and small-molecule inhibitors such as Ispinesib (SB-715992) [74], Filanesib (ARRY-520) [75], and litronesib (LY2523355) [76] have entered clinical trials. However, although these inhibitors have demonstrated excellent efficacies in certain human cancer with no neurotoxicity [75, 77], none have been used as a marketed anticancer agent; thus, further investigation is warranted to in the development of KIF11-based anticancer drugs.

There are several limitations to our study. First, a larger cohort is required to further validate these results. Second,

because of the cautious approach adopted in the study, we failed to enroll the adequate number of HBV-HCC cases with complete clinical characteristics and sufficient long-term follow-up, which did not allow us to conduct informative analyses of better risk stratification and validation. Third, further in-depth studies are necessary to confirm the oncogenic roles of the selected hub genes via *in vitro* and *in vivo* assays.

5. Conclusion

In summary, with the integrated bioinformatics analysis, 309 robust DEGs involved in HBV-HCC were screened, which is helpful for a better understanding of molecular pathogenesis and tumorigenesis of HBV-HCC. Based on a series of comprehensive downstream analysis, 17 potential hub genes were identified that may play critical roles in the development of HBV-HCC. TOP2A and KIF11 can be jointly used to predict overall survival for HBV-HCC, and all the hub genes may hold good potential in exploring early detection biomarkers and therapeutic targets for HBV-HCC.

Data Availability

The data used to support our results are available at the GEO (<https://www.ncbi.nlm.nih.gov/geo/>), TCGA (<https://portal.gdc.cancer.gov/>), and ICGC (<https://icgc.org/>).

Ethical Approval

This article does not contain any studies with human participants or animals performed by any of the authors.

Conflicts of Interest

The authors declare that they have no conflict of interest.

Authors' Contributions

Yuqin Tang and Yongqiang Zhang performed the research and wrote the manuscript. Xun Hu conceived, designed, and supervised the research and provided some useful suggestions. All authors read and approved the final version of the manuscript. Yuqin Tang and Yongqiang Zhang contributed equally to this work.

Acknowledgments

This study was funded by a Research Project of a Pilot Study on Establishing Natural Population Cohorts in Western China (No. 2019YFS0039). We sincerely thank TCGA, ICGC, and GEO databases for their data sharing. In addition, we also acknowledge invaluable help from peer reviewers.

Supplementary Materials

“Supplementary Figure 1: flow chart of the prognostic key gene identification with “multi-split” method. Supplementary Figure 2: validation of 309 overlapping DEGs by integrating analysis with batch removal. Supplementary Figure 3: PPI network of DEGs. Supplementary Figure 4: boxplots showing the relative expression levels of 17 hub genes across

normal liver tissues and cancer tissues with different BCLC stages for HBV-HCC. Supplementary Figure 5: the ROC curves of the selected hub genes to evaluate the diagnostic efficiency for HBV-HCC based on TCGA-LIHC cohort (A-C) and GSE14520 (D-F) cohort. Supplementary Figure 6: stratification analysis of the two-hub gene-based classifier. Supplementary Figure 7: miRNA-mRNA interaction network predicted by miRNet.”. Supplementary Table 1: validation of 309 differentially expressed genes (DEGs) by integrating analysis. Supplementary Table 2: Gene Set Enrichment Analysis (GSEA) result of the two-hub gene-based prognostic signature by the KEGG database. Supplementary Table 3: Gene Set Enrichment Analysis (GSEA) result of the two-hub gene-based prognostic signature by molecular function component of the GO database. (*Supplementary Materials*)

References

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, “Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries,” *CA: a Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.
- [2] M. Zhou, H. Wang, X. Zeng et al., “Mortality, morbidity, and risk factors in China and its provinces, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017,” *Lancet*, vol. 394, no. 10204, pp. 1145–1158, 2019.
- [3] C. de Martel, D. Maucort-Boulch, M. Plummer, and S. Franceschi, “World-wide relative contribution of hepatitis B and C viruses in hepatocellular carcinoma,” *Hepatology*, vol. 62, no. 4, pp. 1190–1200, 2015.
- [4] P. Yang, Q. J. Li, Y. Feng et al., “TGF- β -miR-34a-CCL22 signaling-induced Treg cell recruitment promotes venous metastases of HBV-positive hepatocellular carcinoma,” *Cancer Cell*, vol. 22, no. 3, pp. 291–303, 2012.
- [5] W. W. Feng, O. Wilkins, S. Bang et al., “CD36-mediated metabolic rewiring of breast Cancer cells promotes resistance to HER2-targeted therapies,” *Cell Reports*, vol. 29, no. 11, pp. 3405–3420.e5, 2019.
- [6] A. Diaz-Lagares, J. Mendez-Gonzalez, D. Hervas et al., “A novel epigenetic signature for early diagnosis in lung cancer,” *Clinical Cancer Research*, vol. 22, no. 13, pp. 3361–3371, 2016.
- [7] X. Hao, H. Luo, M. Krawczyk et al., “DNA methylation markers for diagnosis and prognosis of common cancers,” *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 28, pp. 7414–7419, 2017.
- [8] Y. Q. Zhang, W. Y. Wang, J. X. Xue et al., “MicroRNA expression profile on solid subtype of invasive lung adenocarcinoma reveals a panel of four miRNAs to be associated with poor prognosis in Chinese patients,” *Journal of Cancer*, vol. 7, no. 12, pp. 1610–1620, 2016.
- [9] J. Guinney, R. Dienstmann, X. Wang et al., “The consensus molecular subtypes of colorectal cancer,” *Nature Medicine*, vol. 21, no. 11, pp. 1350–1356, 2015.
- [10] W. T. Liu, Y. Wang, J. Zhang et al., “A novel strategy of integrated microarray analysis identifies CENPA, CDK1 and CDC20 as a cluster of diagnostic biomarkers in lung adenocarcinoma,” *Cancer Letters*, vol. 425, pp. 43–53, 2018.
- [11] R. Chen, P. Khatri, P. K. Mazur et al., “A meta-analysis of lung cancer gene expression identifies PTK7 as a survival gene in

- lung adenocarcinoma," *Cancer Research*, vol. 74, no. 10, pp. 2892–2902, 2014.
- [12] Y. Zhang, W. He, and S. Zhang, "Seeking for correlative genes and signaling pathways with bone metastasis from breast cancer by integrated analysis," *Frontiers in Oncology*, vol. 9, 2019.
- [13] Y. Zhou, O. Layton, and L. Hong, "Identification of genes and pathways involved in ovarian epithelial cancer by bioinformatics analysis," *Journal of Cancer*, vol. 9, no. 17, pp. 3016–3022, 2018.
- [14] Q. Ma, Y. Xu, H. Liao et al., "Identification and validation of key genes associated with non-small-cell lung cancer," *Journal of Cellular Physiology*, vol. 234, no. 12, pp. 22742–22752, 2019.
- [15] D. Li, X. Hao, and Y. Song, "An integrated analysis of key microRNAs, regulatory pathways and clinical relevance in bladder cancer," *Oncotargets and Therapy*, vol. 11, pp. 3075–3085, 2018.
- [16] S. Xie, X. Jiang, J. Zhang et al., "Identification of significant gene and pathways involved in HBV-related hepatocellular carcinoma by bioinformatics analysis," *PeerJ*, vol. 7, article e7408, 2019.
- [17] H. Wang, X. Huo, X. R. Yang et al., "STAT3-mediated upregulation of lncRNA HOXD-AS1 as a ceRNA facilitates liver cancer metastasis by regulating SOX4," *Molecular Cancer*, vol. 16, no. 1, 2017.
- [18] I. Ivanovska, C. Zhang, A. M. Liu et al., "Gene signatures derived from a c-MET-driven liver cancer mouse model predict survival of patients with hepatocellular carcinoma," *PLoS One*, vol. 6, no. 9, article e24582, 2011.
- [19] S. Yoo, W. Wang, Q. Wang et al., "A pilot systematic genomic comparison of recurrence risks of hepatitis B virus-associated hepatocellular carcinoma with low- and high-degree liver fibrosis," *BMC Medicine*, vol. 15, no. 1, 2017.
- [20] S. M. Wang, L. L. Ooi, and K. M. Hui, "Identification and validation of a novel gene signature associated with the recurrence of human hepatocellular carcinoma," *Clinical Cancer Research*, vol. 13, no. 21, pp. 6275–6283, 2007.
- [21] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *OMICS*, vol. 16, no. 5, pp. 284–287, 2012.
- [22] C.-H. Chin, S.-H. Chen, H.-H. Wu, C.-W. Ho, M. T. Ko, and C. Y. Lin, "cytoHubba: identifying hub objects and sub-networks from complex interactome," *BMC Systems Biology*, vol. 8, Supplement 4, 2014.
- [23] G. D. Bader and C. W. Hogue, "An automated method for finding molecular complexes in large protein interaction networks," *BMC Bioinformatics*, vol. 4, no. 1, 2003.
- [24] P. Shannon, A. Markiel, O. Ozier et al., "Cytoscape: a software environment for integrated models of biomolecular interaction networks," *Genome research*, vol. 13, no. 11, pp. 2498–2504, 2003.
- [25] R. A. Irizarry, B. Hobbs, F. Collin et al., "Exploration, normalization, and summaries of high density oligonucleotide array probe level data," *Biostatistics*, vol. 4, no. 2, pp. 249–264, 2003.
- [26] Q. Lian, S. Wang, G. Zhang et al., "HCCDB: a database of hepatocellular carcinoma expression atlas," *Genomics, Proteomics & Bioinformatics*, vol. 16, no. 4, pp. 269–275, 2018.
- [27] M. E. Ritchie, B. Phipson, D. Wu et al., "limma powers differential expression analyses for RNA-sequencing and microarray studies," *Nucleic Acids Research*, vol. 43, no. 7, article e47, 2015.
- [28] J. T. Leek, W. E. Johnson, H. S. Parker, A. E. Jaffe, and J. D. Storey, "The sva package for removing batch effects and other unwanted variation in high-throughput experiments," *Bioinformatics*, vol. 28, no. 6, pp. 882–883, 2012.
- [29] A. Lex, N. Gehlenborg, H. Strobel, R. Vuillemot, and H. Pfister, "UpSet: visualization of Intersecting Sets," *IEEE Transactions on Visualization and Computer Graphics*, vol. 20, no. 12, pp. 1983–1992, 2014.
- [30] L. M. Hinder, M. Park, A. E. Rumora et al., "Comparative RNA-Seq transcriptome analyses reveal distinct metabolic pathways in diabetic nerve and kidney disease," *Journal of Cellular and Molecular Medicine*, vol. 21, no. 9, pp. 2140–2152, 2017.
- [31] W. Walter, F. Sánchez-Cabo, and M. Ricote, "GOplot: an R package for visually combining expression data with functional analysis," *Bioinformatics*, vol. 31, no. 17, pp. 2912–2914, 2015.
- [32] X. Robin, N. Turck, A. Hainard et al., "pROC: an open-source package for R and S+ to analyze and compare ROC curves," *BMC Bioinformatics*, vol. 12, no. 1, 2011.
- [33] T. M. Therneau and T. L. Lumley, "Package 'survival,'" *R Top Doc*, vol. 128, 2015.
- [34] Y. Fan, K. Siklenka, S. K. Arora, P. Ribeiro, S. Kimmins, and J. Xia, "miRNet - dissecting miRNA-target interactions and functional associations through network-based visual analysis," *Nucleic Acids Research*, vol. 44, no. W1, pp. W135–W141, 2016.
- [35] M. Liu, Q. Hu, M. Tu et al., "MCM6 promotes metastasis of hepatocellular carcinoma via MEK/ERK pathway and serves as a novel serum biomarker for early recurrence," *Journal of Experimental & Clinical Cancer Research*, vol. 37, no. 1, 2018.
- [36] Q. Yang, B. Xie, H. Tang et al., "Minichromosome maintenance 3 promotes hepatocellular carcinoma radioresistance by activating the NF- κ B pathway," *Journal of Experimental & Clinical Cancer Research*, vol. 38, no. 1, 2019.
- [37] Y. Xiong, J. Lu, Q. Fang et al., "UBE2C functions as a potential oncogene by enhancing cell proliferation, migration, invasion, and drug resistance in hepatocellular carcinoma cells," *Bioscience Reports*, vol. 39, no. 4, 2019.
- [38] R. Satow, M. Shitashige, Y. Kanai et al., "Combined functional genome survey of therapeutic targets for hepatocellular carcinoma," *Clinical Cancer Research*, vol. 16, no. 9, pp. 2518–2528, 2010.
- [39] M. Arai, N. Kondoh, N. Imazeki et al., "The knockdown of endogenous replication factor C4 decreases the growth and enhances the chemosensitivity of hepatocellular carcinoma cells," *Liver International*, vol. 29, no. 1, pp. 55–62, 2009.
- [40] N. Wong, W. Yeo, W. L. Wong et al., "TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance," *International Journal of Cancer*, vol. 124, no. 3, pp. 644–652, 2009.
- [41] M. Sun, G. Wu, Y. Li et al., "Expression profile reveals novel prognostic biomarkers in hepatocellular carcinoma," *Frontiers in Bioscience (Elite Edition)*, vol. 2, pp. 829–840, 2010.
- [42] M. Sun, W. Mo, X. Fu et al., "System biology analysis of cell cycle pathway involved in hepatocellular carcinoma," *Frontiers in Bioscience (Scholar Edition)*, vol. S2, no. 3, pp. 1127–1144, 2010.
- [43] C. X. Wu, X. Q. Wang, S. H. Chok et al., "Blocking CDK1/PDK1/ β -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*, vol. 8, no. 14, pp. 3737–3750, 2018.
- [44] J. Gu, X. Liu, J. Li, and Y. He, "MicroRNA-144 inhibits cell proliferation, migration and invasion in human hepatocellular carcinoma by targeting CCNB1," *Cancer Cell International*, vol. 19, no. 1, 2019.
- [45] S. Shen, J. Kong, Y. Qiu, X. Yang, W. Wang, and L. Yan, "Identification of core genes and outcomes in hepatocellular carcinoma by bioinformatics analysis," *Journal of Cellular Biochemistry*, vol. 120, no. 6, pp. 10069–10081, 2018.
- [46] J. Li, J. Z. Gao, J. Du, Z. X. Huang, and L. X. Wei, "Increased CDC20 expression is associated with development and progression of hepatocellular carcinoma," *International Journal of Oncology*, vol. 45, no. 4, pp. 1547–1555, 2014.
- [47] F. Yang, Y. Hu, H. X. Liu, and Y. J. Wan, "MiR-22-silenced cyclin A expression in colon and liver cancer cells is regulated by bile acid receptor," *The Journal of Biological Chemistry*, vol. 290, no. 10, pp. 6507–6515, 2015.
- [48] S.-J. Fu, J. Chen, F. Ji et al., "MiR-486-5p negatively regulates oncogenic NEK2 in hepatocellular carcinoma," *Oncotarget*, vol. 8, no. 32, article 17635, pp. 52948–52959, 2017.
- [49] L. Y. Huang, C. C. Chang, Y. S. Lee et al., "Inhibition of Hec1 as a novel approach for treatment of primary liver cancer," *Cancer Chemotherapy and Pharmacology*, vol. 74, no. 3, article 2540, pp. 511–520, 2014.
- [50] W. Liao, W. Liu, Q. Yuan et al., "Silencing of DLGAP5 by siRNA significantly inhibits the proliferation and invasion of hepatocellular carcinoma cells," *PLoS One*, vol. 8, no. 12, article e80789, 2013.
- [51] J. Chen, S. Li, S. Zhou et al., "Kinesin superfamily protein expression and its association with progression and prognosis in hepatocellular carcinoma," *Journal of Cancer Research and Therapeutics*, vol. 13, no. 4, pp. 651–659, 2017.
- [52] J.-L. Deng, Y.-H. Xu, and G. Wang, "Identification of potential crucial genes and key pathways in breast cancer using bioinformatic analysis," *Frontiers in Genetics*, vol. 10, 2019.
- [53] R. Zhang, J. Xu, J. Zhao, and J. H. Bai, "Proliferation and invasion of colon cancer cells are suppressed by knockdown of TOP2A," *Journal of Cellular Biochemistry*, vol. 119, no. 9, pp. 7256–7263, 2018.
- [54] J. Shen, S. Yu, X. Sun, M. Yin, J. Fei, and J. Zhou, "Identification of key biomarkers associated with development and prognosis in patients with ovarian carcinoma: evidence from bioinformatic analysis," *Journal of ovarian research*, vol. 12, no. 1, 2019.
- [55] H. Q. Liu, S. L. Zhang, and S. Song, "HER-2/neu and TOP2A expression in gastric cancer reflect disease severity," *Hepato-Gastroenterology*, vol. 59, no. 116, pp. 1290–1293, 2012.
- [56] Z. He, F. Tang, Z. Lu et al., "Analysis of differentially expressed genes, clinical value and biological pathways in prostate cancer," *American Journal of Translational Research*, vol. 10, no. 5, pp. 1444–1456, 2018.
- [57] T. Chen, Y. Sun, P. Ji, S. Kopetz, and W. Zhang, "Topoisomerase II α in chromosome instability and personalized cancer therapy," *Oncogene*, vol. 34, no. 31, pp. 4019–4031, 2015.
- [58] M. Jain, L. Zhang, M. He, Y. Q. Zhang, M. Shen, and E. Kebebew, "TOP2A is overexpressed and is a therapeutic target for adrenocortical carcinoma," *Endocrine-Related Cancer*, vol. 20, no. 3, pp. 361–370, 2013.
- [59] S. Belluti, V. Basile, P. Benatti, E. Ferrari, G. Marverti, and C. Imbriano, "Concurrent inhibition of enzymatic activity and NF-Y-mediated transcription of Topoisomerase-II α by bis-demethoxycurcumin in cancer cells," *Cell Death & Disease*, vol. 4, no. 8, article e756, 2013.
- [60] J. Rapley, M. Nicolas, A. Groen et al., "The NIMA-family kinase Nek6 phosphorylates the kinesin Eg5 at a novel site necessary for mitotic spindle formation," *Journal of Cell Science*, vol. 121, no. 23, pp. 3912–3921, 2008.
- [61] D. J. Sharp, K. R. Yu, J. C. Sisson, W. Sullivan, and J. M. Scholey, "Antagonistic microtubule-sliding motors position mitotic centrosomes in *Drosophila* early embryos," *Nature Cell Biology*, vol. 1, no. 1, pp. 51–54, 1999.
- [62] A. Blangy, H. A. Lane, P. d'Hérin, M. Harper, M. Kress, and E. A. Nigg, "Phosphorylation by p34^{cdc2} regulates spindle association of human Eg5, a kinesin-related motor essential for bipolar spindle formation in vivo," *Cell*, vol. 83, no. 7, pp. 1159–1169, 1995.
- [63] K. M. Bartoli, J. Jakovljevic, J. L. Woolford Jr., and W. S. Saunders, "Kinesin molecular motor Eg5 functions during polypeptide synthesis," *Molecular Biology of the Cell*, vol. 22, no. 18, pp. 3420–3430, 2011.
- [64] J. Zhou, W. R. Chen, L. C. Yang et al., "KIF11 functions as an oncogene and is associated with poor outcomes from breast cancer," *Cancer Research and Treatment*, vol. 51, no. 3, pp. 1207–1221, 2019.
- [65] J. Liu, H. Meng, S. Li et al., "Identification of potential biomarkers in association with progression and prognosis in epithelial ovarian cancer by integrated bioinformatics analysis," *Frontiers in Genetics*, vol. 10, 2019.
- [66] K. Daigo, A. Takano, P. M. Thang et al., "Characterization of KIF11 as a novel prognostic biomarker and therapeutic target for oral cancer," *International Journal of Oncology*, vol. 52, no. 1, pp. 155–165, 2018.
- [67] G. Jungwirth, T. Yu, M. Moustafa et al., "Identification of KIF11 as a novel target in meningioma," *Cancers*, vol. 11, no. 4, 2019.
- [68] E. Terribas, M. Fernández, H. Mazuelas et al., "KIF11 and KIF15 mitotic kinesins are potential therapeutic vulnerabilities for malignant peripheral nerve sheath tumors," *Neuro-Oncology Advances*, vol. 2, Supplement 1, pp. i62–i74, 2020.
- [69] M. A. Schneider, P. Christopoulos, T. Muley et al., "AURKA, DLGAP5, TPX2, KIF11 and CKAP5: five specific mitosis-associated genes correlate with poor prognosis for non-small cell lung cancer patients," *International Journal of Oncology*, vol. 50, no. 2, pp. 365–372, 2017.
- [70] T. U. Mayer, T. M. Kapoor, S. J. Haggarty, R. W. King, S. L. Schreiber, and T. J. Mitchison, "Small molecule inhibitor of mitotic spindle bipolarity identified in a phenotype-based screen," *Science*, vol. 286, no. 5441, pp. 971–974, 1999.
- [71] S. DeBonis, D. A. Skoufias, L. Lebeau et al., "In vitro screening for inhibitors of the human mitotic kinesin Eg5 with antimitotic and antitumor activities," *Molecular Cancer Therapeutics*, vol. 3, no. 9, pp. 1079–1090, 2004.
- [72] S. Hotha, J. C. Yarrow, J. G. Yang et al., "HR22C16: a potent small-molecule probe for the dynamics of cell division," *Angewandte Chemie (International Ed. in English)*, vol. 42, no. 21, pp. 2379–2382, 2003.
- [73] R. Sakowicz, J. T. Finer, C. Beraud et al., "Antitumor activity of a kinesin inhibitor," *Cancer Research*, vol. 64, no. 9, pp. 3276–3280, 2004.

- [74] H. Wakui, N. Yamamoto, S. Kitazono et al., "A phase 1 and dose-finding study of LY2523355 (litronesib), an Eg5 inhibitor, in Japanese patients with advanced solid tumors," *Cancer Chemotherapy and Pharmacology*, vol. 74, no. 1, article 2467, pp. 15–23, 2014.
- [75] J. J. Shah, J. A. Zonder, A. D. Cohen et al., "ARRY-520 shows durable responses in patients with relapsed/refractory multiple myeloma in a phase 1 dose-escalation study," *blood*, vol. 118, no. 21, pp. 1860–1860, 2011.
- [76] S. D. Knight and C. A. Parrish, "Recent progress in the identification and clinical evaluation of inhibitors of the mitotic kinesin KSP," *Current Topics in Medicinal Chemistry*, vol. 8, no. 10, pp. 888–904, 2008.
- [77] C. W. Lee, K. Bélanger, S. C. Rao et al., "A phase II study of ispinesib (SB-715992) in patients with metastatic or recurrent malignant melanoma: a National Cancer Institute of Canada Clinical Trials Group trial," *Investigational New Drugs*, vol. 26, no. 3, article 9097, pp. 249–255, 2008.