Identification of potential hub genes associated with the pathogenesis and prognosis of hepatocellular carcinoma via integrated bioinformatics analysis

Ziqi Meng, Jiarui Wu, Xinkui Liu, Wei Zhou, Mengwei Ni, Shuyu Liu, Siyu Guo, Shanshan Jia and Jingyuan Zhang

Abstract

Objective: The objective was to identify potential hub genes associated with the pathogenesis and prognosis of hepatocellular carcinoma (HCC).

Methods: Gene expression profile datasets were downloaded from the Gene Expression Omnibus database. Differentially expressed genes (DEGs) between HCC and normal samples were identified via an integrated analysis. A protein–protein interaction network was constructed and analyzed using the STRING database and Cytoscape software, and enrichment analyses were carried out through DAVID. Gene Expression Profiling Interactive Analysis and Kaplan–Meier plotter were used to determine expression and prognostic values of hub genes.

Results: We identified 11 hub genes (CDK1, CCNB2, CDC20, CCNB1, TOP2A, CCNA2, MELK, PBK, TPX2, KIF20A, and AURKA) that might be closely related to the pathogenesis and prognosis of HCC. Enrichment analyses indicated that the DEGs were significantly enriched in metabolism-associated pathways, and hub genes and module 1 were highly associated with cell cycle pathway.

Conclusions: In this study, we identified key genes of HCC, which indicated directions for further research into diagnostic and prognostic biomarkers that could facilitate targeted molecular therapy for HCC.

Department of Clinical Chinese Pharmacy, School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing, China

Corresponding author:
Jiarui Wu, Department of Clinical Chinese Pharmacy, School of Chinese Materia Medica, Beijing University of Chinese Medicine, No. 11 North Three-ring East Road, Chaoyang District, Beijing 100029, China.
Email: exogamy@163.com
Introduction

On a global scale, cancer is the main public health problem and liver cancer is a major contributor to both cancer morbidity and mortality.\(^1\) Liver cancer is the sixth most common cancer and the fourth highest cause of cancer-related mortality worldwide.\(^2\) There were expected to be 42,030 newly diagnosed cases and 31,780 deaths of liver cancer in the United States during 2019.\(^3\) Hepatocellular carcinoma (HCC) is the most common form of primary liver cancer, comprising 75% to 85% of cases.\(^2\) The well-recognized risk factors for HCC include chronic infection with hepatitis B (HBV) or hepatitis C virus, exposure to dietary aflatoxin, alcohol-induced cirrhosis, smoking, obesity, and type 2 diabetes.\(^2,4\) In Asia (especially China), chronic HBV infection is the leading etiologic factor of HCC.\(^5\) Most HCC patients are diagnosed at an advanced stage, and locoregional treatments (chemoembolization) and surgical treatments are relatively disappointing in terms of overall survival (OS) of patients with advanced disease.\(^6\) In addition, traditional chemotherapies have not shown promising outcomes in treatment of HCC and have significant toxicity.\(^6,7\) Meanwhile, the lack of early detection of diagnostic markers and limited treatment strategies increase the risk of poor prognosis and death.\(^8\) Therefore, there is a pressing need to develop robust diagnostic strategies and effective therapies for HCC patients.\(^9\)

Over the past decades, microarray technology and bioinformatics have been extensively applied to identify the molecular mechanisms of HCC, which provide strong research support for the diagnosis, treatment, and prognosis of HCC.\(^10\) Because of the ability to process a large number of datasets quickly, integrated bioinformatics analysis and microarray technology have allowed researchers to comprehensively identify the functions of numerous differentially expressed genes (DEGs) in HCC, and they help researchers explore the complicated process of HCC occurrence and development.\(^10,11\) A work by He et al.\(^12\) identified four hub genes and two important pathways in the development of HCC from cirrhosis from one Gene Expression Omnibus (GEO) dataset using a bioinformatics method, including DEG screening, enrichment analyses, and construction of a protein–protein interaction (PPI) network. Zhang et al.\(^13\) screened hub genes and pathways correlated with the occurrence and progression of HCC via a series of bioinformatics analyses incorporating DEGs identification, functional enrichment analyses, PPI network and module analysis, and weighted correlation network analysis. Zhou et al.\(^11\) identified the pivotal genes and microRNAs in HCC using a bioinformatics approach, including analysis of raw data via GEO2R, Gene Ontology (GO), and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses, and construction of PPI network. However, to improve the diagnosis and treatment of HCC, novel diagnostic and prognostic biomarkers for HCC are needed. The flowchart of the study approach is shown in Figure 1.
Figure 1. Flowchart for identification of core genes and pathways for hepatocellular carcinoma (HCC).
GEO, Gene Expression Omnibus; DEG, differentially expressed gene; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, protein–protein interaction; GEPIA, Gene Expression Profiling Interactive Analysis.
Materials and methods

**Ethical approval**

Ethical approval was not required in this study because we analyzed only published data from the GEO database.

**Gene expression profile data**

Gene expression profile data (GSE36376, GSE39791, GSE41804, GSE54236, GSE57957, GSE62232, GSE64041, GSE69715, GSE76427, GSE84005, GSE87630, GSE112790, and GSE121248) were downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/), a public data repository, including high-throughput gene expression and other functional genome datasets. The selection criteria for the included datasets were as follows: (1) tissue samples collected from human HCC and corresponding adjacent or normal tissues; and (2) including at least 40 samples.

**Integrated analysis of microarray datasets**

The matrix data of each GEO dataset were normalized and log2 transformed using the R software package limma, and the DEGs between HCC and corresponding adjacent or normal tissues were also filtered using the limma package. Integration of DEGs identified from the 13 datasets was performed by RobustRankAggreg package in R software. A log2 fold change (FC) ≥1 and adjusted P-value < 0.05 were considered significant for the DEGs.

**Enrichment analyses of DEGs**

Database for Annotation, Visualization and Integrated Discovery (DAVID; https://david.ncifcrf.gov/, version, 6.8) is a comprehensive functional annotation tool for extracting biological significance from large gene/protein datasets. In this study, the GO and KEGG pathway enrichment analyses for the DEGs were conducted via DAVID. The visualization of enrichment analysis results was conducted by using ggplot2 and the GOplot package in the R software.

**PPI network and module analysis**

Search Tool for the Retrieval of Interacting Genes/Proteins (STRING; https://string-db.org/) is a database of known and predicted protein interactions, showing direct and indirect interactions among proteins. This database was applied to obtain potential interactions among the DEGs. PPIs with a confidence score ≥0.7 were reserved and imported into Cytoscape software to construct the PPI network. Furthermore, the clustering modules in this PPI network were analyzed using the MCODE (Molecular Complex Detection) plugin in Cytoscape. Pathway enrichment analyses for important modules were also carried out. The visualization of enrichment analysis results was performed by using the imageGP platform (http://www.ehbio.com/ImageGP/index.php/Home/Index/GOenrichmentplot.html).

**Survival analysis of hub genes**

Kaplan–Meier plotter (KM plotter; http://kmplot.com/analysis/) is a database containing clinical data and gene expression data. This database is used to further understanding the molecular basis of disease and identifying biomarkers associated with survival. The recurrence-free survival and OS information were based on GEO, the European Genome-phenome Archive (EGA), and The Cancer Genome Atlas (TCGA) databases. Hazard ratios (HR) with 95% confidence intervals and log rank P-value were calculated to assess the
association of gene expression with survival and are shown in plots.\textsuperscript{38}

**Expression level analysis and correlation analysis of hub genes**

The Gene Expression Profiling Interactive Analysis (GEPIA; http://gepia.cancer.pku.cn/index.html)\textsuperscript{39} is a newly developed web-based tool that applies a standard processing pipeline to analyze gene expression data between tumor and normal tissues. The relationship of expression of hub genes in HCC and normal tissues were visualized by boxplot.\textsuperscript{38} In addition, correlation analysis was performed by GEPIA to check the relative ratios between two genes.\textsuperscript{39}

### Results

#### Identification of DEGs

In the present study, 13 datasets were downloaded from GEO that included 1100 cancer tissues and 717 corresponding adjacent or normal tissues (Table 1). After integrated analysis, 380 DEGs (293 downregulated and 87 upregulated) were identified (Figure 2a-m and Appendix). Figure 2n shows the top 20 down- and upregulated genes.

#### GO and KEGG pathway enrichment analyses of DEGs

To deepen our understanding of DEGs, we performed GO and KEGG pathway enrichment analyses of DEGs.

**Table 1.** Information for the 13 Gene Expression Omnibus datasets included in the current study.

| Dataset  | Platform                                      | Number of samples (tumor/control) |
|----------|-----------------------------------------------|-----------------------------------|
| GSE36376 | GPL10558-Illumina HumanHT-12 V4.0 expression  | 433 (240/193)                     |
|          | beadchip                                      |                                   |
| GSE39791 | GPL10558-Illumina HumanHT-12 V4.0 expression  | 144 (72/72)                       |
|          | beadchip                                      |                                   |
| GSE41804 | GPL570-[HG-U133_Plus_2] Affymetrix Human      | 40 (20/20)                        |
|          | Genome U133 Plus 2.0 Array                    |                                   |
| GSE54236 | GPL6480-Agilent-014850 Whole Human Genome     | 161 (81/80)                       |
|          | Microarray 4x44K G4112F (Probe Name version)  |                                   |
| GSE57957 | GPL10558-Illumina HumanHT-12 V4.0 expression  | 78 (39/39)                        |
|          | beadchip                                      |                                   |
| GSE62232 | GPL570-[HG-U133_Plus_2] Affymetrix Human      | 91 (81/10)                        |
|          | Genome U133 Plus 2.0 Array                    |                                   |
| GSE64041 | GPL6244-[HuGene-1.0-st] Affymetrix Human Gene 1.0 ST Array [transcript (gene) version]  | 125 (60/65)                       |
|          | Genome U133 Plus 2.0 Array                    |                                   |
| GSE69715 | GPL570-[HG-U133_Plus_2] Affymetrix Human      | 103 (37/66)                       |
|          | Genome U133 Plus 2.0 Array                    |                                   |
| GSE76427 | GPL10558-Illumina HumanHT-12 V4.0 expression  | 167 (115/52)                      |
|          | beadchip                                      |                                   |
| GSE84005 | GPL5175-[HuEx-1.0-st] Affymetrix Human Exon 1.0 ST Array [transcript (gene) version]  | 76 (38/38)                        |
|          | Array                                         |                                   |
| GSE87630 | GPL6947-Illumina HumanHT-12 V3.0 expression   | 94 (64/30)                        |
|          | beadchip                                      |                                   |
| GSE112790| GPL570-[HG-U133_Plus_2] Affymetrix Human      | 198 (183/15)                      |
|          | Genome U133 Plus 2.0 Array                    |                                   |
| GSE121248| GPL570-[HG-U133_Plus_2] Affymetrix Human      | 107 (70/37)                       |
|          | Genome U133 Plus 2.0 Array                    |                                   |
enrichment analyses. Thirty-one significantly enriched GO terms were selected based on a false discovery rate (FDR) < 0.05 (Figure 3a and Appendix). In the GO terms were 13 terms for biological process, mainly related to metabolic process, P450 pathway, and oxidation-reduction process; 12 terms for molecular function, highly involved with multiple enzyme activities, heme binding, iron ion binding and oxygen binding; and 6 terms for cellular components, associated with organelle membrane, extracellular exosome, extracellular region, extracellular space, blood
microparticle, and membrane attack complex.

In the KEGG pathway enrichment analyses, we screened nine pathways according to FDR < 0.05. Figure 3c shows the results of KEGG analysis; the DEGs primarily participated in diverse metabolism-associated signaling pathways, such as metabolic pathways, retinol metabolism, drug metabolism-cytochrome P450, among others.

**PPI network establishment and module analysis**

The PPI network of DEGs comprised 242 nodes and 1267 interactions (Figure 4a); degree was calculated to identify candidate key nodes. Finally, 11 potential key nodes were identified, the degrees of which were all more than four times the corresponding average values: CDK1, CCNB2, CDC20, CCNB1, TOP2A, CCNA2, MELK, PBK, TPX2, KIF20A, and AURKA (Table 2). Moreover, to determine important clustering modules in the PPI network, module analysis was performed using MCODE, and the two modules with the highest scores (score >10) were obtained (Figure 4b, 4c). The enrichment pathways of module 1 and module 2 are shown in Figure 5. Module 1 was highly associated with cell cycle and oocyte meiosis; module 2 was closely connected to drug metabolism-cytochrome P450, linoleic acid metabolism, chemical carcinogenesis, arachidonic acid metabolism, retinol metabolism, metabolism of xenobiotics by cytochrome P450, and metabolic pathways.

**Survival analysis, expression, and correlation analysis of hub genes**

Survival analysis of 11 hub genes was performed using the KM plotter. The results
showed that high expression of **CDK1** (HR = 2.15, 95% CI: 1.52–3.06; \( P = 1.1 \times 10^{-5} \)), **CCNB2** (HR = 1.91, 95% CI: 1.28–2.87; \( P = 0.0013 \)), **CDC20** (HR = 2.49, 95% CI: 1.72–3.59; \( P = 5.1 \times 10^{-7} \)), **CCNB1** (HR = 2.34, 95% CI: 1.55–3.54; \( P = 3.4 \times 10^{-5} \)), **TOP2A** (HR = 1.99, 95% CI: 1.39–2.86; \( P = 0.00012 \)), **CCNA2** (HR = 1.92, 95% CI: 1.36–2.72; \( P = 0.00018 \)), **MELK** (HR = 2.22, 95% CI: 1.5–3.27; \( P = 3.7 \times 10^{-5} \)), **PBK** (HR = 2.24, 95% CI: 1.5–3.34; \( P = 4.8 \times 10^{-5} \)), **TPX2** (HR = 2.29, 95% CI: 1.62–3.24; \( P = 1.4 \times 10^{-6} \)), **KIF20A** (HR = 2.15, 95% CI: 1.52–3.06; \( P = 1.1 \times 10^{-5} \)).

**Figure 4.** PPI network and hub clustering modules. (a) The PPI network of DEGs, (b) module 1 (MCODE score = 38.769), and (c) module 2 (MCODE score = 10.364). Blue circles represent downregulated genes and red circles represent upregulated genes. PPI, protein–protein interaction; DEG, differentially expressed gene; MCODE, Molecular Complex Detection.
(HR = 2.33, 95% CI: 1.63–3.32; P = 1.8e-06), and AURKA (HR = 1.77, 95% CI: 1.25–2.5; P = 0.0011) was related to unfavorable OS for HCC patients (Figure 6). Furthermore, GEPIA was adopted to analyze the different expression level of hub genes in HCC and normal tissues and the 11 hub genes were confirmed to be highly expressed in HCC tissues (Figure 7). The correlations between hub genes were also analyzed by GEPIA, which showed that the 11 hub genes were significantly correlated with each other. Figure 8 showed that the increase in expression of CDK1 was strongly associated with increased expression of the other 10 hub genes.

### Discussion

HCC is the most common type of malignancy and one of the leading causes of cancer-related mortality worldwide. Although much research has been done on HCC, its early diagnosis and treatment remains difficult because of a lack of understanding of the molecular mechanisms associated with HCC occurrence and development. Therefore, in-depth studies of the etiological factors and molecular mechanisms of HCC are of critical importance for HCC diagnosis and treatment. Currently, bioinformatics analysis and microarray technology are developing rapidly and this approach can be used to identify therapeutic targets for diagnosis, therapy, and prognosis of a variety of cancers.

### Table 2. Upregulated hub genes with high degrees.

| Gene  | Degree | Type | MCODE Cluster |
|-------|--------|------|---------------|
| CDK1  | 47     | up   | Module 1      |
| CCNB2 | 46     | up   | Module 1      |
| CDC20 | 45     | up   | Module 1      |
| CCNB1 | 45     | up   | Module 1      |
| TOP2A | 44     | up   | Module 1      |
| CCNA2 | 44     | up   | Module 1      |
| MELK  | 43     | up   | Module 1      |
| PBK   | 43     | up   | Module 1      |
| TPX2  | 43     | up   | Module 1      |
| KIF20A| 43     | up   | Module 1      |
| AURKA | 43     | up   | Module 1      |

Figure 5. Pathway analysis of the two modules with the highest scores. The y-axis shows significantly enriched KEGG pathways, and x-axis shows the two modules. KEGG, Kyoto Encyclopedia of Genes and Genomes; FDR, false discovery rate.
Figure 6. Prognostic roles of 11 hub genes in patients with HCC shown as survival curves. (a) CDK1, (b) CCNB2, (c) CDC20, (d) CCNB1, (e) TOP2A, (f) CCNA2, (g) MELK, (h) PBK, (i) TPX2, (j) KIF20A, and (k) AURKA. HCC, hepatocellular carcinoma; HR, hazard ratio.
Figure 7. Analysis of expression levels of 11 hub genes in human HCC. The red and gray boxes represent cancer and normal tissues, respectively. (a) CDK1, (b) CCNB2, (c) CDC20, (d) CCNB1, (e) TOP2A, (f) CCNA2, (g) MELK, (h) PBK, (i) TPX2, (j) KIF20A, and (k) AURKA. HCC, hepatocellular carcinoma; LIHC, liver hepatocellular carcinoma.
Figure 8. Correlation analysis of 10 hub genes in HCC with CDK1: (a) CCNB2, (b) CDC20, (c) CCNB1, (d) TOP2A, (e) CCNA2, (f) MELK, (g) PBK, (h) TPX2, (i) KIF20A, and (j) AURKA. HCC, hepatocellular carcinoma.
In this research, we identified 380 DEGs, including 293 downregulated and 87 upregulated genes, between HCC and corresponding adjacent or normal tissues. Enrichment analyses indicated that the DEGs were mostly associated with metabolic processes, such as metabolism of retinol, drugs, xenobiotics, tyrosine, tryptophan, and histidine, as well as fatty acid degradation. This indicated that metabolic dysregulation is closely related to HCC. In addition, we obtained 11 hub genes (CDK1, CCNB2, CDC20, CCNB1, TOP2A, CCNA2, MELK, PBK, TPX2, KIF20A, and AURKA) in the PPI network, which were upregulated in HCC tissues compared with normal tissues; expression of the first hub gene, CDK1, was strongly correlated with that of the other hub genes. Overexpression of the 11 hub genes was correlated with worse OS.

Recent evidence implies that tumor cells need specific interphase cyclin-dependent kinases (CDKs) to proliferate. Cyclin-dependent kinase 1 (CDK1) belongs to the CDK family, a member of the serine/threonine protein kinases, and it is crucial for the cell cycle phase transitions G1/S and G2/M. CDK1 is required for cell proliferation because it is the only CDK that can initiate mitosis. The deregulation of CDK1 is likely related to HCC tumorigenesis. Research has found that high expression of CDK1 is correlated with poor OS of HCC. Cyclins act as the regulatory subunits of the CDKs, regulating temporal transitions among various stages of the cell cycle via CDK activation. Cyclin-A2 (CCNA2), cyclin-B1 (CCNB1), and cyclin-B2 (CCNB2), encoded by the CCNA2, CCNB1, and CCNB2 genes, respectively, all belong to the cyclin family. CCNA2 activates CDK1 at the end of interphase to facilitate the onset of mitosis, and CCNA2 overexpression has been reported in numerous types of cancers. A previous study indicated that CCNA2 amplification and overexpression is associated with carcinogenesis of transgenic mouse liver tumors. Moreover, research has demonstrated that inhibition of CCNA2 can arrest HCC cell proliferation and tumorigenesis. High expression of CCNA2 is associated with reduced survival in patients with breast cancer and HCC. CCNB1 and CCNB2 are the principal activators of CDK1 and, together with CDK1, they promote the G2/M transition. Expression of CCNB1 changes periodically throughout the cell cycle, and is a crucial initiator of mitosis. Decreased CCNB1 expression is related to inhibition of HCC occurrence and development, and activation of CCNB1 expression can promote proliferation in human HCC cells. Furthermore, previous research has shown that CCNB1 is closely connected to prognosis of HCC patients. The dimerization of CCNB2 with CDK1 is an essential component of the cell cycle regulatory machinery, and an increase in expression of CCNB2/CDK1 could promote tumor cell proliferation. Furthermore, CCNB2 is highly expressed in several malignant tumors and overexpression of CCNB2 is related to poor prognosis in HCC. Cell division cycle protein 20 (encoded by CDC20) is a regulator of cell cycle checkpoints, which plays a crucial role in anaphase initiation and exit from mitosis. It degrades several important substrates, including cyclin A and CCNB1, to regulate cell cycle progression. CDC20 overexpression is related to progression and poor prognosis in various malignant tumors. Thus, it is a potential target in multiple cancer treatments. A recent study found that increased expression of CDC20 is related to HCC development and progression. Additionally, research has indicated that silencing expression of CDC20 and heparanase can activate cell apoptosis; thus, targeting inhibition of both CDC20 and
heparanase expression is an ideal approach for the treatment of HCC.69

Aurora kinase A (encoded by AURKA) is involved in centrosome duplication, spindle formation, chromosomal amplification and segregation, and cytokinesis, and it plays a significant role in centrosome maturation and mitotic commitment in the late G2 phase.60,71 Abnormal activity of AURKA promotes tumorigenic progression and transformation via defective control at the checkpoint of mitotic spindle.72 Meanwhile, AURKA is highly expressed in a variety of human cancers, including breast cancer,73 lung cancer,74 gastrointestinal cancer,75 bladder cancer,76 and oral cancer.77 A study demonstrated that genetic variations in AURKA might be a reliable predictor of early-stage HCC and a crucial biomarker for HCC development.78 Moreover, other research has indicated that AURKA contributes in metastasis and invasiveness of HCC.79 Therefore, AURKA might represent a new therapeutic target for HCC. Topoisomerase II alpha (TOP2A), a potential biomarker for cancer therapy, has been detected in various types of cancer.80–82 It participates in chromosome condensation and chromatid separation.80 TOP2A encodes topoisomerase II alpha81 and is reported to be overexpressed in HCC tissues.83 Furthermore, a study has shown that TOP2A has prognostic value in HCC and its reactive agents can be used in HCC therapy.84 Maternal embryonic leucine zipper kinase (encoded by MELK) is a member of the AMP protein kinase family of serine/threonine kinases, which affect many stages of tumorigenesis.85 Several studies have shown that MELK is an oncogenic kinase essential for early HCC recurrence, and its expression is upregulated in HCC.86–88 Furthermore, MELK inhibition is associated with suppression of tumor growth, indicating that MELK is a potential therapeutic target for HCC.89 PDZ-binding kinase (encoded by PBK) can regulate cell cycle processes.90 Although PBK is barely detectable in normal somatic tissues, it is often elevated in various tumor tissues and is therefore an important target for cancer screening and targeted therapy.91,92 Recent research has shown that PBK overexpression promotes migration and invasion of HCC, and it could be a therapeutic target for HCC metastasis.93 Targeting protein for Xklp2 (TPX2) expression is modulated by the cell cycle, and it is detected in G1/S phase and disappears after cytokinesis.94,95 Several studies have indicated that TPX2 is highly expressed in different types of cancers.96,97 Additionally, expression of TPX2 is related to proliferation and apoptosis in HCC.98 TPX2 overexpression promotes the growth and metastasis of HCC.99 Kinesin family member 20A (KIF20A) is required during mitosis for the final step of cytokinesis.100,101 Studies have found that high expression of KIF20A is correlated with progression or poor prognosis of many types of cancers.102–104 Furthermore, KIF20A is aberrantly expressed in HCC tissues and its expression may be associated with poor OS.105

According to enrichment analyses of two modules, we found that module 1 was mostly associated with cell cycle and module 2 was closely related to metabolic pathways. Furthermore, all 11 hub genes belonged to module 1 and most are associated with cell cycle and enriched in the “cell cycle” pathway. A number of studies have elucidated that cell cycle disorders are closely related to human cancer.43 Therefore, the carcinogenesis and progression of HCC may be associated with the cell cycle pathway, and we might be able to suppress HCC cell cycle progression, inhibit HCC cell proliferation, and reduce HCC malignancy by downregulating expression of the 11 hub genes identified herein.

Compared with previous studies, this work has several advantages, as follows.
First, the current integrated microarray data used a relatively large sample size from several GEO datasets (GSE36376, GSE39791, GSE41804, GSE54236, GSE57957, GSE62232, GSE64041, GSE69715, GSE76427, GSE84005, GSE87630, GSE112790, and GSE121248). Second, functional enrichment analyses were performed to identify the main biological functions and pathways regulated by DEGs. Third, we established PPI networks, conducted module analysis, discovered potential biomarkers for diagnosis and prognosis of HCC, and performed correlation analysis of hub genes.

The limitations of this work were as follows: First, our results need to be verified by corresponding experimental studies. Second, we obtained data from the GEO database, and data quality cannot be verified. Finally, our study focused on genes that are typically identified as significant changes in diverse datasets, without regard to sex, age, or grading and staging of tumors from which the samples were derived.

**Conclusion**

In the present work, we identified 11 hub genes (CDK1, CCNB2, CDC20, CCNB1, TOP2A, CCNA2, MELK, PBK, TPX2, KIF20A, and AURKA) associated with the development and poor prognosis of HCC by integrated bioinformatics analysis. However, because our study was based on data analysis only, further experiments are required to confirm the results. Our findings will provide evidence and new insights to enhance approaches for the early diagnosis, prognosis, and treatment of HCC.

**Declaration of conflicting interest**

The authors declare that there is no conflict of interest.

**Funding**

This work was supported by the National Natural Science Foundation of China (No. 81473547, 81673829) and the Young Scientists Training Program of Beijing University of Chinese Medicine.

**ORCID iD**

Jiarui Wu https://orcid.org/0000-0002-1617-6110

**References**

1. Feng RM, Zong YN, Cao SM, et al. Current cancer situation in China: good or bad news from the 2018 Global Cancer Statistics? Cancer Commun (Lond) 2019; 39: 22. DOI: 10.1186/s40880-019-0368-6.

2. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018; 68: 394–424. DOI: 10.3322/caac.21492.

3. Siegel RL, Miller KD and Jemal A. Cancer statistics, 2019. CA Cancer J Clin 2019; 69: 7–34. DOI: 10.3322/caac.21551.

4. Toh TB, Lim JJ and Chow EK. Epigenetics of hepatocellular carcinoma. Clin Transl Med 2019; 8: 13. DOI: 10.1186/s40169-019-0230-0.

5. Ogunwobi OO, Harricharran T, Huaman J, et al. Mechanisms of hepatocellular carcinoma progression. World J Gastroenterol 2019; 25: 2279–2293. DOI: 10.3748/wjg.v25.i19.2279.

6. Chen S, Cao Q, Wen W, et al. Targeted therapy for hepatocellular carcinoma: challenges and opportunities. Cancer Lett 2019; 460: 1–9. DOI: 10.1016/j.canlet.2019.114428.

7. Eatrides J, Wang E, Kothari N, et al. Role of systemic therapy and future directions for hepatocellular carcinoma. Cancer Control 2017; 24: 1073274817729243. DOI: 10.1177/1073274817729243.

8. Wang X, Liao Z, Bai Z, et al. MiR-93-5p promotes cell proliferation through down-regulating PPARGC1A in hepatocellular carcinoma cells by bioinformatics analysis and experimental verification. Genes
9. Li C, Zhou D, Jiang X, et al. Identifying hepatocellular carcinoma-related hub genes by bioinformatics analysis and CYP2C8 is a potential prognostic biomarker. *Gene* 2019; 698: 9–18. DOI: 10.1016/j.gene.2019.02.062.

10. Liu S, Yao X, Zhang D, et al. Analysis of transcription factor-related regulatory networks based on bioinformatics analysis and validation in hepatocellular carcinoma. *Biomed Res Int* 2018; 2018: 1431396. DOI: 10.1155/2018/1431396.

11. Zhou L, Du Y, Kong L, et al. Identification of molecular target genes and key pathways in hepatocellular carcinoma by bioinformatics analysis. *Onco Targets Ther* 2018; 11: 1861–1869. DOI: 10.2147/OTT.S156737.

12. He B, Yin J, Gong S, et al. Bioinformatics analysis of key genes and pathways for hepatocellular carcinoma transformed from cirrhosis. *Medicine (Baltimore)* 2017; 96: e6938. DOI: 10.1097/MD.0000000000006938.

13. Zhang C, Peng L, Zhang Y, et al. The identification of key genes and pathways in hepatocellular carcinoma by bioinformatics analysis of high-throughput data. *Med Oncol* 2017; 34: 101. DOI: 10.1007/s12032-017-0963-9.

14. Lim HY, Sohn I, Deng S, et al. Prediction of disease-free survival in hepatocellular carcinoma by gene expression profiling. *Ann Surg Oncol* 2013; 20: 3747–3753. DOI: 10.1245/s10434-013-3070-y.

15. Kim JH, Sohn BH, Lee HS, et al. Genomic predictors for recurrence patterns of hepatocellular carcinoma: model derivation and validation. *PLoS Med* 2014; 11: e1001770. DOI: 10.1371/journal.pmed.1001770.

16. Hodo Y, Honda M, Tanaka A, et al. Association of interleukin-28B genotype and hepatocellular carcinoma recurrence in patients with chronic hepatitis C. *Clin Cancer Res* 2013; 19: 1827–1837. DOI: 10.1158/1078-0432.CCR-12-1641.

17. Villa E, Critelli R, Lei B, et al. Neoangiogenesis-related genes are hallmarks of fast-growing hepatocellular carcinomas and worst survival. Results from a prospective study. *Gut* 2016; 65: 861–869. DOI: 10.1136/gutjnl-2014-308483.

18. Zubiete-Franco I, Garcia-Rodriguez JL, Lopitz-Otsoa F, et al. SUMOylation regulates LKB1 localization and its oncogenic activity in liver cancer. *EBioMedicine* 2019; 40: 406–421. DOI: 10.1016/j.ebiom.2018.12.031.

19. Mah WC, Thurnherr T, Chow PK, et al. Methylation profiles reveal distinct subgroup of hepatocellular carcinoma patients with poor prognosis. *PLoS One* 2014; 9: e104158. DOI: 10.1371/journal.pone.0104158.

20. Schulze K, Imbeaud S, Letouze E, et al. Exome sequencing of hepatocellular carcinomas identifies new mutational signatures and potential therapeutic targets. *Nat Genet* 2015; 47: 505–511. DOI: 10.1038/ng.3252.

21. Makowska Z, Boldanova T, Adametz D, et al. Gene expression analysis of biopsy samples reveals critical limitations of transcriptome-based molecular classifications of hepatocellular carcinoma. *J Pathol Clin Res* 2016; 2: 80–92. DOI: 10.1002/cjp2.37.

22. Sekhar V, Pollicino T, Diaz G, et al. Infection with hepatitis C virus depends on TACSTD2, a regulator of Claudin-1 and occludin highly downregulated in hepatocellular carcinoma. *PLoS Pathog* 2018; 14: e1006916. DOI: 10.1371/journal.ppat.1006916.

23. Grinchuk OV, Yenamandra SP, Iyer R, et al. Tumor-adjacent tissue co-expression profile analysis reveals pro-oncogenic ribosomal gene signature for prognosis of resectable hepatocellular carcinoma. *Mol Oncol* 2018; 12: 89–113. DOI: 10.1002/1878-0261.12153.

24. Woo HG, Choi JH, Yoon S, et al. Integrative analysis of genomic and epigenomic regulation of the transcriptome in liver cancer. *Nat Commun* 2017; 8: 839. DOI: 10.1038/s41467-017-00991-w.

25. Shimada S, Mogushi K, Akiyama Y, et al. Comprehensive molecular and immunological characterization of hepatocellular
carcinoma. *EBioMedicine* 2019; 40: 457–470. DOI: 10.1016/j.ebiom.2018.12.058.

26. Wang SM, Ooi LL and Hui KM. Identification and validation of a novel gene signature associated with the recurrence of human hepatocellular carcinoma. *Clin Cancer Res* 2007; 13: 6275–6283. DOI: 10.1158/1078-0432.CCR-06-2236.

27. Clough E and Barrett T. The Gene Expression Omnibus database. *Methods Mol Biol* 2016; 1418: 93–110. DOI: 10.1007/978-1-4939-3578-9_5.

28. Ritchie ME, Phipson B, Wu D, et al. *limma* powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43: e47. DOI: 10.1093/nar/gkv007.

29. Kolde R, Laur S, Adler P, et al. Robust rank aggregation for gene list integration and meta-analysis. *Bioinformatics* 2012; 28: 573–580. DOI: 10.1093/bioinformatics/btr709.

30. Huang da W, Sherman BT and Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 2009; 4: 44–57. DOI: 10.1038/nprot.2008.211.

31. Ito K and Murphy D. Application of ggplot2 to pharmacometric graphics. *CPT Pharmacometrics Syst Pharmacol* 2013; 2: e79. DOI: 10.1038/psp.2013.56.

32. Walter W, Sanchez-Cabo F and Ricote M. GOpplot: an R package for visually combining expression data with functional analysis. *Bioinformatics* 2015; 31: 2912–2914. DOI: 10.1093/bioinformatics/btv300.

33. Szklarczyk D, Morris JH, Cook H, et al. The STRING database in 2017: quality-controlled protein-protein association networks, made broadly accessible. *Nucleic Acids Res* 2017; 45: D362–D368. DOI: 10.1093/nar/gkw937.

34. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13: 2498–2504. DOI: 10.1101/gr.1239303.

35. Bader GD and Hogue CW. An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics* 2003; 4: 2.

36. Hou GX, Liu P, Yang J, et al. Mining expression and prognosis of topoisomerase isoforms in non-small-cell lung cancer by using Oncomine and Kaplan-Meier plotter. *PLoS One* 2017; 12: e0174515. DOI: 10.1371/journal.pone.0174515.

37. Lacny S, Wilson T, Clement F, et al. Kaplan-Meier survival analysis overestimates cumulative incidence of health-related events in competing risk settings: a meta-analysis. *J Clin Epidemiol* 2018; 93: 25–35. DOI: 10.1016/j.jclinepi.2017.10.006.

38. Sun C, Yuan Q, Wu D, et al. Identification of core genes and outcome in gastric cancer using bioinformatics analysis. *Oncotarget* 2017; 8: 70271–70280. DOI: 10.18632/oncotarget.20082.

39. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. *Nucleic Acids Res* 2017; 45: W98–W102. DOI: 10.1093/nar/gkx247.

40. Hage C, Hoves S, Ashoff M, et al. Characterizing responsive and refractory orthotopic mouse models of hepatocellular carcinoma in cancer immunotherapy. *PLoS One* 2017; 14: e0219517. DOI: 10.1371/journal.pone.0219517.

41. Mou T, Zhu D, Wei X, et al. Identification and interaction analysis of key genes and microRNAs in hepatocellular carcinoma by bioinformatics analysis. *World J Surg Oncol* 2017; 15: 63. DOI: 10.1186/s12957-017-1127-2.

42. Xu Z, Zhou Y, Cao Y, et al. Identification of candidate biomarkers and analysis of prognostic values in ovarian cancer by integrated bioinformatics analysis. *Med Oncol* 2016; 33: 130. DOI: 10.1007/s12032-016-0840-y.

43. Malumbres M and Barbacid M. Cell cycle, CDKs and cancer: a changing paradigm. *Nat Rev Cancer* 2009; 9: 153–166. DOI: 10.1038/nrc2602.

44. Doree M and Hunt T. From Cdc2 to Cdk1: when did the cell cycle kinase join its cyclin partner? *J Cell Sci* 2002; 115: 2461–2464.

45. Wu CX, Wang XQ, Chok SH, et al. 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 2018; 8: 3737–3750. DOI: 10.7150/thno.25487.

46. Bednarek K, Kiwerska K, Szaumkessel M, et al. Recurrent CDK1 overexpression in laryngeal squamous cell carcinoma. Tumour Biol 2016; 37: 11115–11126. DOI: 10.1007/s13277-016-4991-4.

47. Zhao J, Han SX, Ma JL, et al. The role of CDK1 in apoptosis induced by TMA in hepatocellular carcinoma cells. Oncol Rep 2013; 30: 253–259. DOI: 10.3892/or.2013.2426.

48. Kim DH, Park SE, Kim M, et al. A functional single nucleotide polymorphism at the promoter region of cyclin A2 is associated with increased risk of colon, liver, and lung cancers. Cancer 2011; 117: 4080–4091. DOI: 10.1002/cncr.25930.

49. Kabir MF, Mohd Ali J and Haji Hashim O. Microarray gene expression profiling in colorectal (HCT116) and hepatocellular (HepG2) carcinoma cell lines treated with Melicope ptelefolia leaf extract reveals transcriptome profiles exhibiting anticancer activity. PeerJ 2018; 6: e5203. DOI: 10.7717/peerj.5203.

50. Haddad R, Morrow AD, Plass C, et al. Restriction landmark genomic scanning of mouse liver tumors for gene amplification: overexpression of cyclin A2. Biochem Biophys Res Commun 2000; 274: 188–196. DOI: 10.1006/bbrc.2000.3124.

51. Yang F, Gong J, Wang G, et al. Waltonite inhibits proliferation of hepatoma cells and tumorigenesis via FXR-miR-22-CCNA2 signaling pathway. Oncotarget 2016; 7: 75165–75175. DOI: 10.18632/oncotarget.12614.

52. Bukholm IR, Bukholm G and Nesland JM. Over-expression of cyclin A is highly associated with early relapse and reduced survival in patients with primary breast carcinomas. Int J Cancer 2001; 93: 283–287.

53. Zhang L, Huang Y, Ling J, et al. Screening and function analysis of hub genes and pathways in hepatocellular carcinoma via bioinformatics approaches. Cancer Biomark 2018; 22: 511–521. DOI: 10.3233/CBM-171160.

54. Huang Y, Sramkoski RM and Jacobberger JW. The kinetics of G2 and M transitions regulated by B cyclins. PLoS One 2013; 8: e80861. DOI: 10.1371/journal.pone.0080861.

55. Gao CL, Wang GW, Yang GQ, et al. Karyopherin subunit-alpha 2 expression accelerates cell cycle progression by upregulating CCNB2 and CDK1 in hepatocellular carcinoma. Oncol Lett 2018; 15: 2815–2820. DOI: 10.3892/ol.2017.7691.

56. Gu J, Liu X, Li J, et al. MicroRNA-144 inhibits cell proliferation, migration and invasion in human hepatocellular carcinoma by targeting CCNB1. Cancer Cell Int 2019; 19: 15. DOI: 10.1186/s12935-019-0729-x.

57. Chai N, Xie HH, Yin JP, et al. FOXM1 promotes proliferation in human hepatocellular carcinoma cells by transcriptional activation of CCNB1. Biochem Biophys Res Commun 2018; 500: 924–929. DOI: 10.1016/j.bbrc.2018.04.201.

58. Weng L, Du J, Zhou Q, et al. Identification of cyclin B1 and Sec62 as biomarkers for recurrence in patients with HBV-related hepatocellular carcinoma after surgical resection. Mol Cancer 2012; 11: 39. DOI: 10.1186/1476-4598-11-39.

59. Li R, Jiang X, Zhang Y, et al. Cyclin B2 overexpression in human hepatocellular carcinoma is associated with poor prognosis. Arch Med Res 2019; 50: 10–17. DOI: 10.1016/j.arcmed.2019.03.003.

60. Yu H. Cdc20: a WD40 activator for a cell cycle degradation machine. Mol Cell 2007; 27: 3–16. DOI: 10.1016/j.molcel.2007.06.009.

61. Fang G, Yu H and Kirschner MW. Direct binding of CDC20 protein family members activates the anaphase-promoting complex in mitosis and G1. Mol Cell 1998; 2: 163–171.

62. Geley S, Kramer E, Gieffers C, et al. Anaphase-promoting complex/cyclosome-dependent proteolysis of human cyclin A starts at the beginning of mitosis and is not subject to the spindle assembly
63. Shirayama M, Toth A, Galova M, et al. APC(Cdc20) promotes exit from mitosis by destroying the anaphase inhibitor Pds1 and cyclin Clb5. Nature 1999; 402: 203–207. DOI: 10.1038/46080.

64. Zhang Q, Huang H, Liu A, et al. Cell division cycle 20 (CDC20) drives prostate cancer progression via stabilization of beta-catenin in cancer stem-like cells. EBioMedicine 2019; 42: 397–407. DOI: 10.1016/j.ebiom.2019.03.032.

65. Cheng S, Castillo V and Sliva D. CDC20 associated with cancer metastasis and novel mushroom-derived CDC20 inhibitors with antimitastatic activity. Int J Oncol 2019; 54: 2250–2256. DOI: 10.3892/ijo.2019.4791.

66. Kato T, Daigo Y, Aragaki M, et al. Overexpression of CDC20 predicts poor prognosis in primary non-small cell lung cancer patients. J Surg Oncol 2012; 106: 423–430. DOI: 10.1002/jso.23109.

67. Li J, Gao JZ, Du JL, et al. Increased CDC20 expression is associated with development and progression of hepatocellular carcinoma. Int J Oncol 2014; 45: 1547–1555. DOI: 10.3892/ijo.2014.2559.

68. Wang L, Zhang J, Wan L, et al. Targeting Cdc20 as a novel cancer therapeutic strategy. Pharmacol Ther 2015; 151: 141–151. DOI: 10.1016/j.pharmthera.2015.04.002.

69. Liu M, Zhang Y, Liao Y, et al. Evaluation of the antitumor efficacy of RNAi-mediated inhibition of CDC20 and heparanase in an orthotopic liver tumor model. Cancer Biother Radiopharm 2015; 30: 233–239. DOI: 10.1089/cbr.2014.1799.

70. Huang CH, Chen CJ, Chen PN, et al. Impacts of AURKA genetic polymorphism on urothelial cell carcinoma development. J Cancer 2019; 10: 1370–1374. DOI: 10.7150/jca.30014.

71. Su ZL, Su CW, Huang YL, et al. A novel AURKA mutant-induced early-onset severe hepatocarcinogenesis greater than wild-type via activating different pathways in zebrafish. Cancers (Basel) 2019; 11: pii: E927. DOI: 10.3390/cancers11070927.

72. Guo M, Lu S, Huang H, et al. Increased AURKA promotes cell proliferation and predicts poor prognosis in bladder cancer. BMC Syst Biol 2018; 12: 118. DOI: 10.1186/s12918-018-0634-2.

73. Cox DG, Hankinson SE and Hunter DJ. Polymorphisms of the AURKA (STK15/Aurora Kinase) gene and breast cancer risk (United States). Cancer Causes Control 2006; 17: 81–83. DOI: 10.1007/s10552-005-0429-9.

74. Lo Iacono M, Monica V, Saviozzi S, et al. Aurora Kinase A expression is associated with lung cancer histological-subtypes and with tumor de-differentiation. J Transl Med 2011; 9: 100. DOI: 10.1186/1479-5876-9-100.

75. Wang-Bishop L, Chen Z, Gomaa A, et al. Inhibition of AURKA reduces proliferation and survival of gastrointestinal cancer cells with activated KRAS by preventing activation of RPS6KB1. Gastroenterology 2019; 156: 662–675.e7. DOI: 10.1053/j.gastro.2018.10.030.

76. Mobley A, Zhang S, Bondaruk J, et al. Aurora kinase A is a biomarker for bladder cancer detection and contributes to its aggressive behavior. Sci Rep 2017; 7: 40714. DOI: 10.1038/srep40714.

77. Huang C, Wang L, Song H, et al. Interactive effects of AURKA polymorphisms with smoking on the susceptibility of oral cancer. Artif Cells Nanomed Biotechnol 2019; 47: 2333–2337. DOI: 10.1080/21691401.2019.1601101.

78. Wang B, Hsu CJ, Chou CH, et al. Variations in the AURKA gene: biomarkers for the development and progression of hepatocellular carcinoma. Int J Med Sci 2018; 15: 170–175. DOI: 10.7150/ijms.22513.

79. Chen C, Song G, Xiang J, et al. AURKA promotes cancer metastasis by regulating epithelial-mesenchymal transition and cancer stem cell properties in hepatocellular carcinoma. Biochem Biophys Res Commun 2017; 486: 514–520. DOI: 10.1016/j.bbrc.2017.03.075.

80. Song J, Ma Q, Hu M, et al. The inhibition of miR-144-3p on cell proliferation and metastasis by targeting TOP2A in HCMV-positive glioblastoma cells.
Molecules 2018; 23: pii: E3259. DOI: 10.3390/molecules23123259.
81. Ghisoni E, Maggiorotto F, Borella F, et al. TOP2A as marker of response to pegylated liposomal doxorubicin (PLD) in epithelial ovarian cancers. *J Ovarian Res* 2019; 12: 17. DOI: 10.1186/s13048-019-0492-6.
82. Zhang R, Xu J, Zhao J, et al. Proliferation and invasion of colon cancer cells are suppressed by knockdown of TOP2A. *J Cell Biochem* 2018; 119: 7256–7263. DOI: 10.1002/jcb.26916.
83. Panvichian R, Tantiwetrueangdet A, Angkathunyakul N, et al. TOP2A amplification and overexpression in hepatocellular carcinoma tissues. *Biomed Res Int* 2015; 2015: 381602. DOI: 10.1155/2015/381602.
84. Wong N, Yeo W, Wong WL, et al. TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance. *Int J Cancer* 2009; 124: 644–652. DOI: 10.1002/ijc.23968.
85. Jiang P and Zhang D. Maternal embryonic leucine zipper kinase (MELK): a novel regulator in cell cycle control, embryonic development, and cancer. *Int J Mol Sci* 2013; 14: 21551–21560. DOI: 10.3390/ijms141121551.
86. Xia H, Kong SN, Chen J, et al. MELK is an oncogenic kinase essential for early hepatocellular carcinoma recurrence. *Cancer Lett* 2016; 383: 85–93. DOI: 10.1016/j.canlet.2016.09.017.
87. Li Y, Li Y, Chen Y, et al. MicroRNA-214-3p inhibits proliferation and cell cycle progression by targeting MELK in hepatocellular carcinoma and correlates cancer prognosis. *Cancer Cell Int* 2017; 17: 102. DOI: 10.1186/s12935-017-0471-1.
88. Hiwatashi K, Ueno S, Sakoda M, et al. Expression of Maternal Embryonic Leucine Zipper Kinase (MELK) correlates to malignant potentials in hepatocellular carcinoma. *Anticancer Res* 2016; 36: 5183–5188. DOI: 10.21873/anticanceres.11088.
89. Chlenski A, Park C, Dobratic M, et al. Maternal Embryonic Leucine Zipper Kinase (MELK), a potential therapeutic target for neuroblastoma. *Mol Cancer Ther* 2019; 18: 507–516. DOI: 10.1158/1535-7163.MCT-18-0819.
90. Su TC, Chen CY, Tsai WC, et al. Cytoplasmic, nuclear, and total PBK/TOPK expression is associated with prognosis in colorectal cancer patients: a retrospective analysis based on immunohistochemistry stain of tissue microarrays. *PLoS One* 2018; 13: e0204866. DOI: 10.1371/journal.pone.0204866.
91. Zhang Y, Yang X, Wang R, et al. Prognostic value of PDZ-binding kinase/T-LAK cell- originated protein kinase (PBK/TOPK) in patients with cancer. *J Cancer* 2019; 10: 131–137. DOI: 10.7150/jca.28216.
92. Ma H, Li Y, Wang X, et al. PBK, targeted by EVI1, promotes metastasis and confers cisplatin resistance through inducing autophagy in high-grade serous ovarian carcinoma. *Cell Death Dis* 2019; 10: 166. DOI: 10.1038/s41419-019-1415-6.
93. Yang QX, Zhong S, He L, et al. PBK overexpression promotes metastasis of hepatocellular carcinoma via activating ETV4-uPAR signaling pathway. *Cancer Lett* 2019; 452: 90–102. DOI: 10.1016/j.canlet.2019.03.028.
94. Zou J, Huang RY, Jiang FN, et al. Overexpression of TPX2 is associated with progression and prognosis of prostate cancer. *Oncol Lett* 2018; 16: 2823–2832. DOI: 10.3892/ol.2018.9016.
95. Zou Z, Zheng B, Li J, et al. TPX2 level correlates with cholangiocarcinoma cell proliferation, apoptosis, and EMT. *Biomed Pharmacother* 2018; 107: 1286–1293. DOI: 10.1016/j.biopha.2018.08.011.
96. Chen M, Zhang H, Zhang G, et al. Targeting TPX2 suppresses proliferation and promotes apoptosis via repression of the PI3k/AKT/P21 signaling pathway and activation of p53 pathway in breast cancer. *Biochem Biophys Res Commun* 2018; 507: 74–82. DOI: 10.1016/j.bbrc.2018.10.164.
97. Aguirre-Portoles C, Bird AW, Hyman A, et al. Tpx2 controls spindle integrity, genome stability, and tumor development. *Cancer Res* 2012; 72: 1518–1528. DOI: 10.1158/0008-5472.CAN-11-1971.
98. Liang B, Jia C, Huang Y, et al. TPX2 level correlates with hepatocellular carcinoma cell proliferation, apoptosis, and EMT. *Dig Dis Sci* 2015; 60: 2360–2372. DOI: 10.1007/s10620-015-3730-9.

99. Huang Y, Guo W and Kan H. TPX2 is a prognostic marker and contributes to growth and metastasis of human hepatocellular carcinoma. *Int J Mol Sci* 2014; 15: 18148–18161. DOI: 10.3390/ijms151018148.

100. Shen T, Yang L, Zhang Z, et al. KIF20A affects the prognosis of bladder cancer by promoting the proliferation and metastasis of bladder cancer cells. *Dis Markers* 2019; 2019: 4863182. DOI: 10.1155/2019/4863182.

101. Morita H, Matsuoka A, Kida JI, et al. KIF20A, highly expressed in immature hematopoietic cells, supports the growth of HL60 cell line. *Int J Hematol* 2018; 108: 607–614. DOI: 10.1007/s12185-018-2527-y.

102. Taniuchi K, Furihata M and Saibara T. KIF20A-mediated RNA granule transport system promotes the invasiveness of pancreatic cancer cells. *Neoplasia* 2014; 16: 1082–1093. DOI: 10.1016/j.neo.2014.10.007.

103. Duan J, Huang W and Shi H. Positive expression of KIF20A indicates poor prognosis of glioma patients. *Onco Targets Ther* 2016; 9: 6741–6749. DOI: 10.2147/OTT.S115974.

104. Zhang W, He W, Shi Y, et al. High expression of KIF20A is associated with poor overall survival and tumor progression in early-stage cervical squamous cell carcinoma. *PLoS One* 2016; 11: e0167449. DOI: 10.1371/journal.pone.0167449.

105. Lu M, Huang X, Chen Y, et al. Aberrant KIF20A expression might independently predict poor overall survival and recurrence-free survival of hepatocellular carcinoma. *IUBMB Life* 2018; 70: 328–335. DOI: 10.1002/iub.1726.
Appendix

Information for 293 downregulated genes (down) and 87 upregulated genes (up).

| Name       | logFC  | Type | Name       | logFC  | Type | Name       | logFC  | Type |
|------------|--------|------|------------|--------|------|------------|--------|------|
| CLEC1B     | -3.33713 | down | IL13RA2    | -1.41685 | down | CSRNP1     | -1.20759 | down |
| C9         | -2.93972 | down | PAMR1      | -1.30729 | down | ZGPAT      | -1.283655 | down |
| FCN3       | -3.32589 | down | CYP26A1    | -1.82557 | down | FAM150B    | -1.096361 | down |
| CYP1A2     | -3.61576 | down | JCHAIN     | -1.90133 | down | LPA        | -1.568535 | down |
| HAMP       | -3.72675 | down | ADIRF      | -1.34189 | down | ALPL       | -1.135143 | down |
| SLC21B3    | -2.84405 | down | NNMT       | -1.65555 | down | S100A8     | -1.149369 | down |
| SPP2       | -2.19210 | down | TAT        | -1.77239 | down | GPM6A      | -1.287388 | down |
| APOF       | -2.7681  | down | MS4A6A     | -1.02381 | down | RCL1       | -1.112209 | down |
| NAT2       | -2.42415 | down | VNN1       | -1.43431 | down | CYP2B7P    | -1.31568 | down |
| CLRN3      | -2.35658 | down | HSD17B2    | -1.27883 | down | CCBE1      | -1.131678 | down |
| RDH16      | -2.05491 | down | FAM134B    | -1.27241 | down | LINC01093  | -1.711116 | down |
| SLC25A47   | -2.3928  | down | CTH        | -1.2995 | down  | ST3GAL6    | -1.008844 | down |
| SLC22A1    | -2.49578 | down | ACACI      | -1.06823 | down | TBX15      | -1.105089 | down |
| THRSP      | -2.37999 | down | OTC        | -1.12724 | down | BCO2       | -1.572843 | down |
| CLEC4G     | -2.8104  | down | CYP2A7     | -1.7189  | down | LUM        | -1.123456 | down |
| GBA3       | -2.26827 | down | C6         | -1.48624 | down | ESR1       | -1.022446 | down |
| DNASE1L3   | -2.22313 | down | GREM2      | -1.17719 | down | CYR6I      | -1.101151 | down |
| SHBG       | -1.96811 | down | HPD        | -1.56635 | down | HBA2       | -1.227362 | down |
| LY6E       | -2.01561 | down | KBTBD11    | -1.69651 | down | KDM8       | -1.06201 | down |
| CDHR2      | -2.02873 | down | CA2        | -1.30707 | down | GADD45G    | -1.126764 | down |
| TMEM27     | -2.33949 | down | AKR7A3     | -1.25278 | down | ASPG       | -1.055061 | down |
| C7         | -2.2597  | down | RNF125     | -1.03098 | down | FCR2B      | -1.141195 | down |
| FBP1       | -1.79884 | down | TTC36      | -1.69649 | down | ASPA       | -1.025006 | down |
| SRD5A2     | -1.89056 | down | PROM1      | -1.44661 | down | PBLD       | -1.006234 | down |
| MT1M       | -3.02758 | down | ADH6       | -1.22168 | down | HHIP       | -1.37843 | down |
| BBOX1      | -2.04999 | down | ETNPP1     | -1.15368 | down | CRP        | -1.053533 | down |
| APOA5      | -1.774  | down | HSD17B13   | -1.50866 | down | FREM2      | -1.522232 | down |
| IGFBP3     | -1.70456 | down | ANXA10     | -1.62516 | down | ADRA1A     | -1.161964 | down |
| ADH4       | -2.15911 | down | FXYD1      | -1.41243 | down | CNTN3      | -1.176196 | down |
| KMO        | -1.91086 | down | GGDH       | -1.30838 | down | ITLN1      | -1.034492 | down |
| CYPB1      | -1.76864 | down | PON1       | -1.17061 | down | UGT2B10    | -1.031179 | down |
| CXL14      | -2.31161 | down | ACSM3      | -1.52866 | down | DIRAS3     | -1.123875 | down |
| GHR        | -2.12511 | down | SLC27A5    | -1.33347 | down | STEAP4     | -1.061309 | down |
| ADGRG7     | -1.85853 | down | LIFR       | -1.47372 | down | CYP4A22    | -1.074568 | down |
| MARCO      | -2.25079 | down | HABP2      | -1.06311 | down | TFPI2      | -1.00071 | down |
| MT1F       | -2.59948 | down | GRAMD1C    | -1.07675 | down | MT1A       | -1.093671 | down |
| CYP39A1     | -1.86139 | down | TKFC       | -1.07859 | down | RAB25      | -1.081375 | down |
| OIT3       | -2.4803  | down | STEAP3     | -1.09586 | down | RDH5       | -1.006888 | down |
| MBL2       | -1.62953 | down | IL1RAP     | -1.21549 | down | EPCAM      | -1.336797 | down |
| VIPR1      | -1.89347 | down | GCDH       | -1.02343 | down | SPINK1     | 3.633978 | up |
| TDO2       | -1.44452 | down | HAL        | -1.262  | down | GPC3       | 2.807155 | up |
| BHMT       | -1.68706 | down | GABARAP1   | -1.07919 | down | AKR1B10    | 2.588879 | up |

(continued)
### Appendix. Continued.

| Name             | logFC   | Type   | Name             | logFC   | Type   | Name             | logFC   |
|------------------|---------|--------|------------------|---------|--------|------------------|---------|
| PKC1             | -1.85362 | down   | ID1              | -1.32236 | down   | ASPM             | 1.804629 |
| MTH1             | -2.20509 | down   | INMT             | -1.65209 | down   | CAP2             | 2.086341 |
| AFM              | -1.90272 | down   | SKAP1            | -1.06342 | down   | TOP2A            | 2.232845 |
| HGFA             | -2.18902 | down   | FETUB            | -1.31249 | down   | PRCI             | 1.923672 |
| MT1G             | -2.64319 | down   | CFHR4            | -1.07478 | down   | CDKN3            | 1.778974 |
| CYP2A6           | -2.05548 | down   | HSD11B1          | -1.27605 | down   | CDC20            | 1.910919 |
| CETP             | -1.77384 | down   | G6PC             | -1.00804 | down   | PTTG1            | 1.451774 |
| SMIM24           | -1.81333 | down   | MFAP4            | -1.53268 | down   | NCAPI            | 1.551838 |
| FCN2             | -1.90705 | down   | ABCA8            | -1.10284 | down   | LN2              | 1.551605 |
| FOSB             | -2.12121 | down   | CYP2J2           | -1.03103 | down   | CCL20            | 1.667526 |
| ECM1             | -1.72876 | down   | AKR1D1           | -1.77452 | down   | FAM83D           | 1.570755 |
| MT1X             | -2.07498 | down   | GPD1             | -1.01057 | down   | KIF20A           | 1.644679 |
| SLC10A1          | -1.70131 | down   | HAO1             | -1.0889  | down   | PBK              | 1.6372  |
| CRHBP            | -2.55698 | down   | TACSTD2          | -1.09909 | down   | AURKA            | 1.321582 |
| F9               | -1.86997 | down   | GCGR             | -1.51767 | down   | UBE2T            | 1.429052 |
| SRPX             | -1.99247 | down   | C8orf4           | -1.53773 | down   | NUSAP1           | 1.447842 |
| CYP2C9           | -1.7781  | down   | DMGDH            | -1.11277 | down   | AKR1C3           | 1.315793 |
| GNMT             | -1.80416 | down   | PON3             | -1.07722 | down   | MELK             | 1.397481 |
| CYP2CB           | -1.84304 | down   | MAT1A            | -1.15605 | down   | SRXN1            | 1.101781 |
| PGLYR2P2         | -1.57039 | down   | AADAT            | -1.45288 | down   | HMMR             | 1.429779 |
| LECT2            | -1.71324 | down   | HPX              | -1.1201  | down   | COL15A1          | 1.679907 |
| HAO2             | -2.05962 | down   | KCNN2            | -1.76035 | down   | UBD              | 1.793116 |
| FOS              | -2.10062 | down   | ACADL            | -1.62191 | down   | PLVAP            | 1.303945 |
| ANGPTL6          | -1.40198 | down   | SLC13A5          | -1.18455 | down   | HSPB1            | 1.057592 |
| CNDP1            | -2.19859 | down   | ASS1             | -1.22714 | down   | SPP1             | 1.372928 |
| CXCL12           | -1.91941 | down   | PRSS8            | -1.15745 | down   | CENPF            | 1.339564 |
| AGXT2            | -1.39193 | down   | CPED1            | -1.24941 | down   | SQLE             | 1.28364 |
| ACOT12           | -1.27878 | down   | FTCD             | -1.25547 | down   | CEPS5            | 1.130246 |
| RSPO3            | -1.62341 | down   | TEMEM45A         | -1.37559 | down   | KIF4A            | 1.431933 |
| PZP              | -1.76877 | down   | ALDH6A1          | -1.08996 | down   | TRIP13           | 1.223148 |
| COLEC10          | -1.85319 | down   | SLC27A2          | -1.02491 | down   | S100P            | 1.428178 |
| HOGA1            | -1.43807 | down   | ETFDH            | -1.15312 | down   | DLGAP5           | 1.462148 |
| MT1E             | -1.80442 | down   | GCKR             | -1.00475 | down   | ALDH3A1          | 1.048498 |
| CYP3A4           | -2.39818 | down   | OAT              | -1.35234 | down   | CDC25            | 1.222277 |
| SLC39A5          | -1.47867 | down   | SFRPS            | -1.04433 | down   | SFN              | 1.002947 |
| KLKB1            | -1.57229 | down   | CYP3A43          | -1.2044  | down   | ESM1             | 1.15394 |
| LCAT             | -1.87391 | down   | SLC6A12          | -1.11241 | down   | TTK              | 1.378481 |
| IGFALS           | -1.94508 | down   | SOCS2            | -1.38986 | down   | TPX2             | 1.091732 |
| GLYAT            | -1.72131 | down   | CYP4F2           | -1.0376  | down   | PAGE4            | 1.240802 |
| ADH1C            | -1.64914 | down   | PHYHD1           | -1.0017  | down   | COL4A1           | 1.236208 |
| PROZ             | -1.52487 | down   | SLC7A2           | -1.05182 | down   | HJRUP            | 1.034534 |
| CYP2E1           | -2.04247 | down   | C1R              | -1.01827 | down   | RAGC1P           | 1.407851 |
| GSTZ1            | -1.39923 | down   | PLG              | -1.09969 | down   | IGF2BP3          | 1.019851 |
| CHST4            | -1.72521 | down   | CPS1             | -1.29626 | down   | ANLN             | 1.53779 |
| MFSD2A           | -1.51912 | down   | ADAMTSL2         | -1.24169 | down   | MCM2             | 1.109517 |
| IDO2             | -1.83679 | down   | MTTP             | -1.02368 | down   | UBE2C            | 1.0809  |

(continued)
| Name    | logFC  | Type  | Name    | logFC  | Type  | Name    | logFC  | Type  |
|---------|--------|-------|---------|--------|-------|---------|--------|-------|
| SDS     | -1.75694 | down  | CXCL2   | -1.43349 | down  | NQO1    | 1.365462 | up    |
| ENO3    | -1.37195 | down  | HRG     | -1.00696 | down  | CCNB2   | 1.303069 | up    |
| GLS2    | -1.75439 | down  | ACSL1   | -1.14524 | down  | CCNA2   | 1.185444 | up    |
| DCN     | -1.94676 | down  | MAN1C1  | -1.18965 | down  | MUC13   | 1.14796  | up    |
| PLAC8   | -1.80012 | down  | PCOLCE  | -1.00609 | down  | MCM6    | 1.016314 | up    |
| SERPINA4| -1.2352  | down  | MT2A    | -1.54319 | down  | CENPW   | 1.083208 | up    |
| ZG16    | -1.56869 | down  | CD1D    | -1.02692 | down  | TGM3    | 0.050965 | up    |
| BCHE    | -1.77407 | down  | XDH     | -1.11927 | down  | RADS1API| 0.049223 | up    |
| CFP     | -1.47416 | down  | PPP1R1A | -1.10299 | down  | THY1    | 0.104685 | up    |
| SLC38A4 | -1.32606 | down  | HBB     | -1.31952 | down  | NUF2    | 1.25884  | up    |
| ADH1A   | -1.27277 | down  | RBP5    | -1.04885 | down  | CKAP2L  | 1.054397 | up    |
| CLEC4M  | -2.35545 | down  | CFHR3   | -1.10107 | down  | MAGEA1  | 1.282995 | up    |
| CYP4A11 | -1.5036  | down  | RELN    | -1.02856 | down  | ECT2    | 1.065576 | up    |
| GYS2    | -1.66608 | down  | NPY1R   | -1.34248 | down  | ACSL4   | 0.16679  | up    |
| PHGDH   | -1.40019 | down  | CLDN10  | -1.34641 | down  | MDK     | 0.076885 | up    |
| BGN     | -1.2236  | down  | ATF5    | -1.11652 | down  | PEG10   | 1.104051 | up    |
| CIDEB   | -1.27052 | down  | GNE     | -1.04957 | down  | COX7B2  | 1.333566 | up    |
| CYP2C19 | -1.55814 | down  | CYP4V2  | -1.05634 | down  | CCNB1   | 1.362239 | up    |
| IYD     | -1.22582 | down  | CDSL    | -1.49237 | down  | RRM2    | 1.542665 | up    |
| C8A     | -1.49471 | down  | TIMD4   | -1.24178 | down  | REG3A   | 1.140254 | up    |
| STAB2   | -1.82665 | down  | EGR1    | -1.41173 | down  | CDK1    | 1.236442 | up    |
| CDA     | -1.14527 | down  | GADD45B | -1.21416 | down  | KIF14   | 1.054151 | up    |
| HPGD    | -1.37821 | down  | GPT2    | -1.15763 | down  | ZIC2    | 1.320155 | up    |
| OLHML3  | -1.38115 | down  | ACMSD   | -1.02364 | down  | BUB1B   | 1.118801 | up    |
| PTH1R   | -1.35746 | down  | CCL19   | -1.32425 | down  | NDC80   | 1.234218 | up    |
| EPHX2   | -1.29488 | down  | RBPI    | -1.15142 | down  | NEK2    | 1.144213 | up    |
| COLEC11 | -1.34767 | down  | ACADS   | -1.05741 | down  | RBM24   | 1.220962 | up    |
| CYP2C18 | -1.121134 | down | MYOM2  | -1.03989 | down  | NMRA1P1 | 1.314053 | up    |
| AMDH1I  | -1.14346 | down  | DCXR    | -1.01852 | down  | DTL     | 1.283296 | up    |
| LYVE1   | -1.69466 | down  | PLGBl   | -1.07364 | down  | SULT1C2 | 1.181554 | up    |
| GSPT2   | -1.16851 | down  | CYP2B6  | -1.37318 | down  | ROBO1   | 1.247873 | up    |
| C8B     | -1.16715 | down  | UROCI   | -1.06129 | down  | SSX1    | 0.001365 | up    |
| ADH1B   | -1.77846 | down  | PDK4    | -1.08546 | down  | FLVCR1  | 0.006476 | up    |
| DPT     | -1.68413 | down  | PPARGC1A | -1.08395 | down  | CTHR1C | 1.120384 | up    |
| AZGP1   | -1.23501 | down  | NDRG2   | -0.01145 | down  | ZWINT   | 0.066653 | up    |
| ALDH8A1 | -1.37768 | down  | IGFI    | -1.14785 | down  | GINS    | 0.103249 | up    |
| RND3    | -1.62821 | down  | ASPDH   | -1.15589 | down  | SMPX    | 1.089408 | up    |
| SLC19A3 | -1.18742 | down  | DBH     | -1.50296 | down  | GPR158  | 0.061576 | up    |
| WDR72   | -1.27875 | down  | PRG4    | -1.13337 | down  |       |        |       |

FC, fold change.
Information on Gene Ontology (GO) enrichment analysis in biological process (BP), cellular component (CC), and molecular function (MF) categories.

| Category | ID       | Term                                                                 | \(-\log_{10}(FDR)\) | Count |
|----------|----------|----------------------------------------------------------------------|-----------------------|-------|
| BP       | GO:0055114 | Oxidation—reduction process                                        | 16.45646128           | 56    |
| BP       | GO:0019373 | Epoxigenase P450 pathway                                             | 12.72414085           | 13    |
| BP       | GO:0006805 | Xenobiotic metabolic process                                        | 6.801196269           | 16    |
| BP       | GO:0017144 | Drug metabolic process                                              | 6.713310124           | 11    |
| BP       | GO:0045926 | Negative regulation of growth                                       | 5.354060264           | 9     |
| BP       | GO:0071276 | Cellular response to cadmium ion                                    | 4.258416753           | 8     |
| BP       | GO:0042738 | Exogenous drug catabolic process                                    | 3.873727759           | 7     |
| BP       | GO:0071294 | Cellular response to zinc ion                                       | 3.86110044             | 8     |
| BP       | GO:0008202 | Steroid metabolic process                                           | 3.349012692           | 10    |
| BP       | GO:0097267 | Omega—hydroxylase P450 pathway                                       | 3.048831706           | 6     |
| BP       | GO:0016098 | Monoterpenoid metabolic process                                     | 2.284734835           | 5     |
| BP       | GO:0007067 | Mitotic nuclear division                                            | 1.901221899           | 19    |
| BP       | GO:0006569 | Tryptophan catabolic process                                        | 1.382839511           | 5     |
| CC       | GO:0031090 | Organelle membrane                                                  | 12.13504583           | 21    |
| CC       | GO:0070062 | Extracellular exosome                                               | 10.96203625           | 117   |
| CC       | GO:0005576 | Extracellular region                                                 | 8.944226201           | 78    |
| CC       | GO:0005615 | Extracellular space                                                 | 8.079401711           | 68    |
| CC       | GO:0072562 | Blood microparticle                                                 | 3.941029653           | 17    |
| CC       | GO:0005579 | Membrane attack complex                                             | 2.131478756           | 5     |
| MF       | GO:0016705 | Oxidoreductase activity, acting on paired donors,                    | 12.77849851           | 19    |
|          |          | with incorporation or reduction of molecular oxygen                 |                       |       |
| MF       | GO:0020037 | Heme binding                                                        | 11.82105086           | 25    |
| MF       | GO:0004497 | Monoxygenase activity                                                | 11.5463498            | 18    |
| MF       | GO:0005506 | Iron ion binding                                                    | 10.69763162           | 25    |
| MF       | GO:0008392 | Arachidonic acid epoxygenase activity                                | 10.22404973           | 11    |
| MF       | GO:0019825 | Oxygen binding                                                       | 9.168975245           | 15    |
| MF       | GO:0016491 | Oxidoreductase activity                                             | 5.664542324           | 22    |
| MF       | GO:0008395 | Steroid hydroxylase activity                                        | 5.613513145           | 10    |
| MF       | GO:0070330 | aromatase activity                                                  | 2.805232257           | 8     |
| MF       | GO:0004024 | Alcohol dehydrogenase activity, zinc—dependent                       | 2.38141982            | 5     |
| MF       | GO:0016712 | Oxidoreductase activity, acting on paired donors,                     | 1.824019096           | 6     |
|          |          | with incorporation or reduction of molecular oxygen, reduced flavin  |                       |       |
|          |          | or flavoprotein as one donor, and incorporation of one atom of       |                       |       |
|          |          | oxygen                                                              |                       |       |
| MF       | GO:0004745 | Retinol dehydrogenase activity                                      | 1.391280368           | 6     |

FDR, false discovery rate.