

Original Article

Identification of key genes and carcinogenic pathways in hepatitis B virus-associated hepatocellular carcinoma through bioinformatics analysis

Sang-Hoon Kim¹, Shin Hwang¹, Gi-Won Song¹, Dong-Hwan Jung¹, Deok-Bog Moon¹, Jae Do Yang², Hee Chul Yu²

¹Division of Liver Transplantation and Hepatobiliary Surgery, Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, Seoul, Korea,

²Department of Surgery, Jeonbuk National University Hospital, Jeonju, Korea

Backgrounds/Aims: Mechanisms for the development of hepatocellular carcinoma (HCC) in hepatitis B virus (HBV)-infected patients remain unclear. The aim of the present study was to identify genes and pathways involved in the development of HBV-associated HCC.

Methods: The GSE121248 gene dataset, which included 70 HCCs and 37 adjacent liver tissues, was downloaded from the Gene Expression Omnibus database. Differentially expressed genes (DEGs) in HCCs and adjacent liver tissues were identified. Gene ontology and Kyoto Encyclopedia of Genes and Genome pathway enrichment analyses were then performed.

Results: Of 134 DEGs identified, 34 were up-regulated and 100 were down-regulated in HCCs. The 34 up-regulated DEGs were mainly involved in nuclear division, organelle fission, spindle and midbody formation, histone kinase activity, and p53 signaling pathway, whereas the 100 down-regulated DEGs were involved in steroid and hormone metabolism, collagen-coated extracellular matrix, oxidoreductase activity, and activity on paired donors, including incorporation or reduction of molecular oxygen, monooxygenase activity, and retinol metabolism. Analyses of protein-protein interaction networks with a high degree of connectivity identified significant modules containing 14 hub genes, including *ANLN*, *ASPM*, *BUB1B*, *CCNB1*, *CDK1*, *CDKN3*, *ECT2*, *HMMR*, *NEK2*, *PBK*, *PRC1*, *RACGAP1*, *RRM2*, and *TOP2A*, which were mainly associated with nuclear division, organelle fission, spindle formation, protein serine/threonine kinase activity, p53 signaling pathway, and cell cycle.

Conclusions: This study identified key genes and carcinogenic pathways that play essential roles in the development of HBV-associated HCC. This may provide important information for the development of diagnostic and therapeutic targets for HCC.

Key Words: Carcinogenesis; Genes; Hepatitis B virus; Bioinformatics; Oncogenesis

INTRODUCTION

Hepatocellular carcinoma (HCC), which represents about 75% to 85% of primary liver cancers, is the sixth most com-

monly diagnosed cancer and the fourth leading cause of cancer-related deaths worldwide. Most HCCs develop in patients with liver cirrhosis, with hepatitis B virus (HBV) infection being the leading cause of liver cancer and liver cancer deaths globally (33%) [1]. Although multiple genes and molecular pathways were shown to be involved in the development of HCC, the exact mechanisms underlying HCC tumorigenesis and progression remain unclear. Identification of key genes and molecular mechanisms underlying the development of HCC is important for determining future diagnostic biomarkers and therapeutic targets.

High-throughput sequencing methods such as microarray analysis can enable the profiling of large gene expression datasets for various diseases including HCC. Thus, it may provide

Received: July 16, 2021, **Revised:** August 30, 2021,
Accepted: September 2, 2021

Corresponding author: Shin Hwang
Department of Surgery, Asan Medical Center, University of Ulsan College of Medicine, 88 Olympic-ro 43-gil, Songpa-gu, Seoul 05505, Korea
Tel: +82-2-3010-3930, Fax: +82-2-3010-6701, E-mail: shwang@amc.seoul.kr
ORCID: <https://orcid.org/0000-0002-9045-2531>



Copyright © The Korean Association of Hepato-Biliary-Pancreatic Surgery
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

clues for understanding the mechanism of pathogenesis including carcinogenesis. Combinations of bioinformatics and microarray technology have identified various key genes essential for the development of HCC [2,3].

The present study analyzed the GSE121248 gene dataset of 70 HCCs and 37 adjacent normal liver tissue samples from patients with HBV-associated HCC obtained from the Gene Expression Omnibus (GEO) database. This dataset was investigated to identify differentially expressed genes (DEGs) in HCCs and adjacent non-cancerous liver tissue samples. Using bioinformatics methods, key genes and pathways were identified using a series of functional analyses.

MATERIALS AND METHODS

Microarray data

The GSE121248 gene dataset based on the GPL570 platform was downloaded from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>). This dataset included RNA array data for 107 tissue samples obtained from patients with HBV-associated HCC, including 70 HCCs and 37 adjacent normal liver tissue samples.

Identification of differentially expressed genes

Data normalization and differential expression analyses were performed using R packages *affy* and *limma* (R Foundation for Statistical Computing, Vienna, Austria) and the GEO2R online tool (<http://www.ncbi.nlm.nih.gov/geo/geo2r/>). Probe IDs without matches or multiple gene matches were removed. Adjusted *p*-values were applied to correct false-positive results using the Benjamini and Hochberg false discovery rate method. Adjusted *p*-values < 0.05 and absolute $\log_2 FC$ values > 2 were considered cut-off criteria for defining DEGs.

Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genome (KEGG) pathway enrichment analysis

To explore the biological significance of DEGs, the Database for Annotation, Visualization and Integrated Discovery (DAVID, <http://david.ncifcrf.gov>) and R packages *clusterProfiler* were used for visualization and enrichment analyses, including analyses of biological processes, cellular components, molecular functions, and KEGG pathways. A *p*-value < 0.05 was considered statistically significant.

Integration of protein-protein interaction network

To determine connections between proteins encoded by DEGs and significant gene modules involved in the development of HCC in HBV-infected patients, protein-protein interaction (PPI) network analysis was performed using the Search Tool for the Retrieval of Interacting Genes (STRING) database (<http://www.string-db.org/>). Cytoscape (3.8.0) was applied to construct a PPI network with a confidence score ≥ 0.4 . Modules were analyzed with Cytoscape plugins Molecular Complex

Detection (MCODE) and GOClue. A *p*-value < 0.05 was considered statistically significant.

Validation of hub genes

To ensure the credibility of microarray analysis of GSE121248, expression levels of mRNAs encoded by 14 hub genes were compared in the GSE121248 and The Cancer Genome Atlas (TCGA) databases. The Gene Expression Profiling Interactive Analysis (GEPIA; <http://gepia.cancer-pku.cn/index.html>), a web server for analyzing the expression of mRNA sequences in tumor and normal tissue samples from the TCGA and Genotype-Tissue Expression (GTEx) projects, was utilized to construct boxplots to visualize the expression of mRNAs encoded by hub DEGs in HCC and normal liver tissue samples based on the TCGA database.

Overall patient survival

The GEPIA database was utilized to construct Kaplan–Meier curves for the overall survival of patients with HCC as a function of the expression of the 14 hub genes. The median value of gene expression was used as the cutoff value for grouping. Curves were compared by log-rank tests. Hazard ratios and 95% confidence intervals were calculated. A *p*-value < 0.05 was considered statistically significant. The study protocol was approved by the Institutional Review Board of Asan Medical Center (IRB No. 2021-0683), which waived the requirement for informed consent due the use of publicly open data.

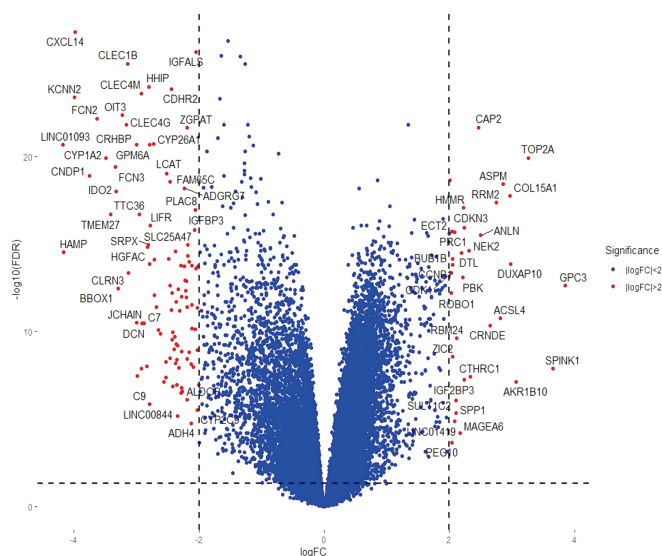


Fig. 1. Volcano plot of differentially expressed genes identified in hepatocellular carcinoma (HCC) vs. adjacent normal liver tissues obtained from patients with hepatitis B virus-associated HCC. FDR, false discovery rate; FC, fold change.

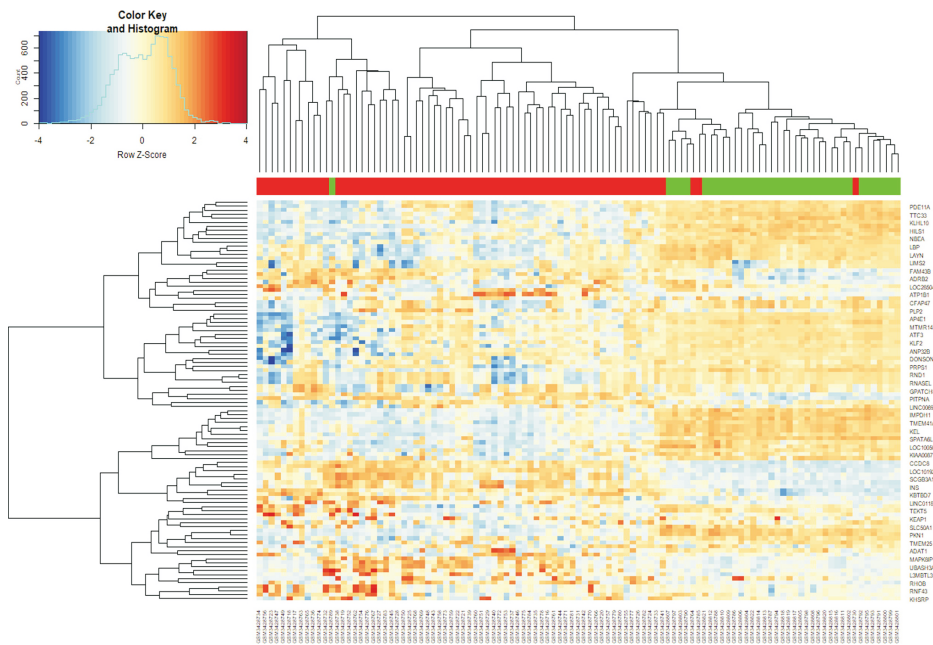


Fig. 2. Heatmap of the top 100 differentially expressed genes (DEGs) with $\log_2 FC > 2$. Blue, down-regulated DEGs; red, up-regulated DEGs. FC, fold change.

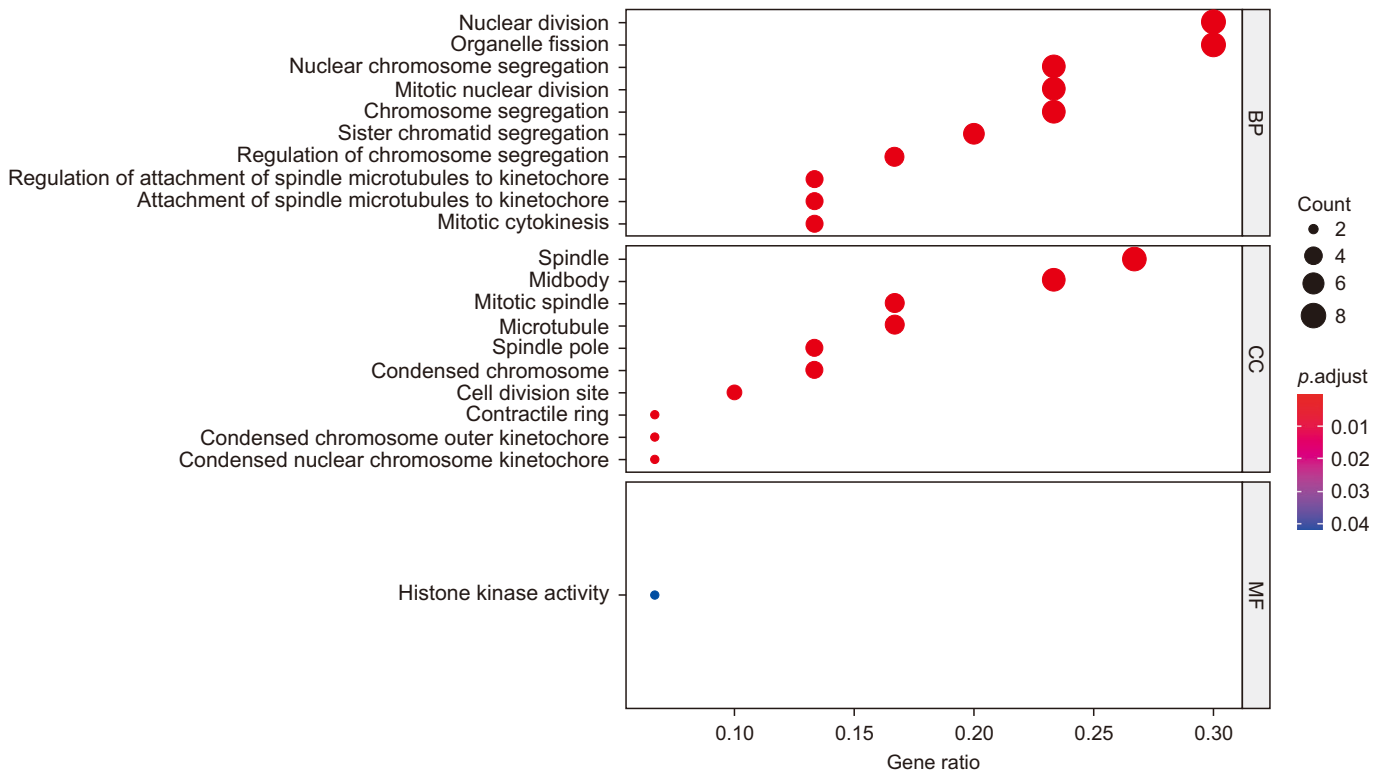


Fig. 3. Gene ontology enrichment analysis of the 34 up-regulated differentially expressed genes with $\log_2 FC > 2$. FC, fold change; BP, biological process; CC, cellular component; MF, molecular function.

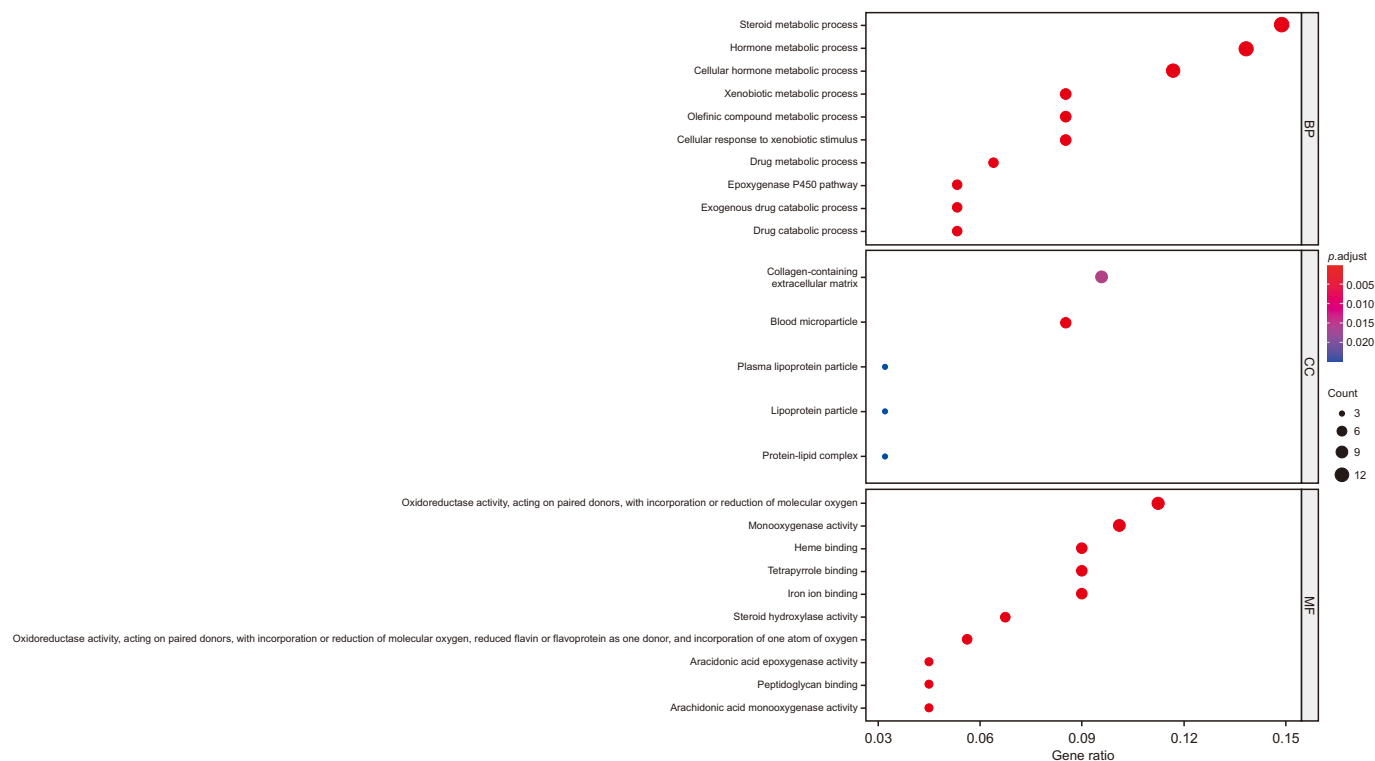


Fig. 4. Gene ontology enrichment analysis of the 100 down-regulated differentially expressed genes with $\log_2 FC < -2$. FC, fold change; BP, biological process; CC, cellular component; MF, molecular function.

RESULTS

Identification of DEGs

The GSE121248 dataset, which included 70 HCC tissue samples and 37 adjacent normal tissue samples from patients with HBV-associated HCC, was analyzed using the R packages *affy* and *limma* and the GEO2R online tool to identify DEGs with thresholds of $|\log_2 FC| > 2$ and adjusted p -values < 0.05 . Of 134 DEGs identified, 34 were significantly higher and 100 significantly lower in HCC than in adjacent normal liver tissues. Fig. 1 is a volcano plot showing the distribution of identified DEGs and Fig. 2 is a heatmap showing expression levels of the 100 top DEGs with $|\log_2 FC| > 2$.

GO and KEGG pathway enrichment analysis

To determine the function of the identified DEGs, functional and pathway enrichment analyses were performed using DAVID and R packages *clusterProfiler*. GO enrichment analysis for assessing biological processes revealed that the 34 up-regulated DEGs were mainly involved in nuclear division and organelle fission (Fig. 3), whereas the 100 down-regulated DEGs were mainly involved in steroid and hormone metabolism (Fig. 4). Analysis of cellular components showed that up-regulated DEGs mainly participated in spindle and midbody formation (Fig. 3), whereas down-regulated DEGs mainly participated

in the synthesis of collagen-containing extracellular matrices and blood microparticles (Fig. 4). Analysis of molecular function showed that up-regulated DEGs were enriched in histone kinase activity (Fig. 3), whereas down-regulated DEGs were

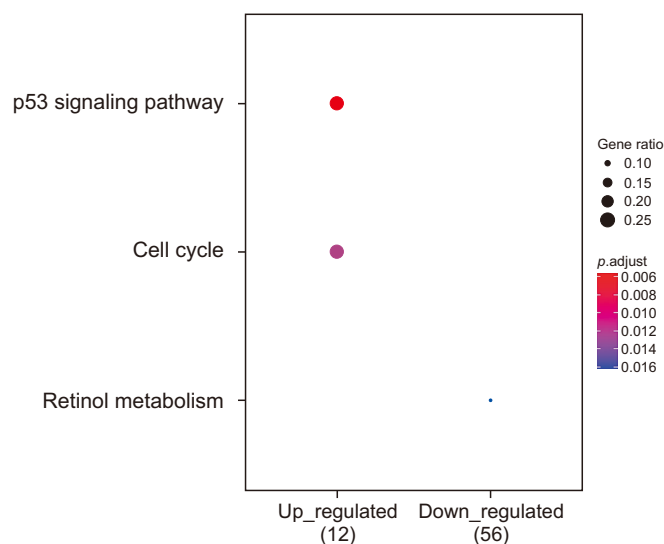


Fig. 5. Enriched Kyoto Encyclopedia of Genes and Genome pathway analysis of the 134 differentially expressed genes.

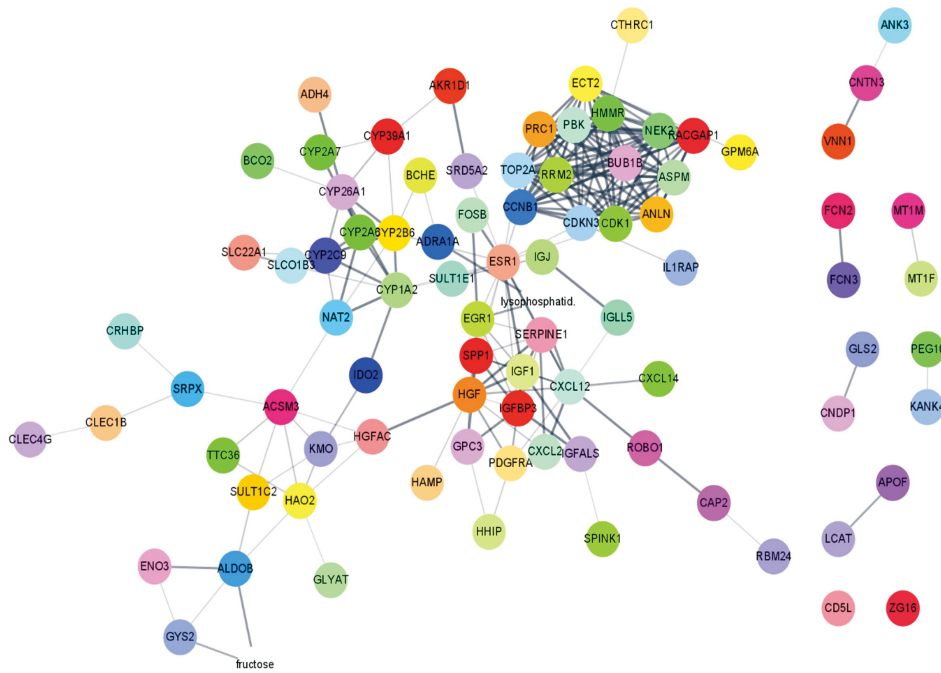


Fig. 6. Overall protein-protein interaction network.

mainly enriched in oxidoreductase activity and activities related to paired donors, including the incorporation or reduction of molecular oxygen and monooxygenase activity (Fig. 4). KEGG pathway analysis found that up-regulated DEGs were mainly involved in the p53 signaling pathway and the cell cycle, whereas down-regulated DEGs were mainly involved in retinol metabolism (Fig. 5).

Integration of the protein-protein interaction network

As constructed by Cytoscape (3.8.0), based on the STRING online tool, the PPI network of the 134 identified DEGs was found to consist of 134 nodes and 370 edges, with an average node degree of 5.82 and an average local clustering coefficient of 0.525 (Fig. 6). Based on the degree of connectivity of these 134 DEGs, the Cytoscape plug-in MCODE was able to create a significant module that included 14 nodes and 91 edges (Fig. 7). These 14 nodes represent 14 hub genes of high connectivity, including *ANLN*, *ASPM*, *BUB1B*, *CCNB1*, *CDK1*, *CDKN3*, *ECT2*, *HMMR*, *NEK2*, *PBK*, *PRC1*, *RACGAP1*, *RRM2*, and *TOP2A*. Subsequent enrichment analysis showed that these 14 hub genes were mainly associated with nuclear division, organelle fission, spindle and midbody formation, protein serine/threonine kinase activity, histone kinase activity, the p53 signaling pathway, and the cell cycle (Fig. 8, Table 1).

Validation of hub genes by the TCGA database

Analysis of expression levels of mRNAs encoded by the 14 hub genes in the GSE121248 dataset showed that all 14 were expressed at higher levels in the 70 HCC samples than in the 34 adjacent normal tissue samples (Fig. 9). To validate these

findings, mRNA expression levels of the 14 hub genes in HCC and normal samples in the TCGA database and GEPIA database were determined (Fig. 10). Similar to findings in the GSE121248 dataset, levels of mRNAs encoded by all 14 hub genes were significantly higher in HCC than in normal liver tissues in the GEPIA database.

Overall patient survival

Kaplan–Meier curves obtained from GEPIA (Fig. 11) showed

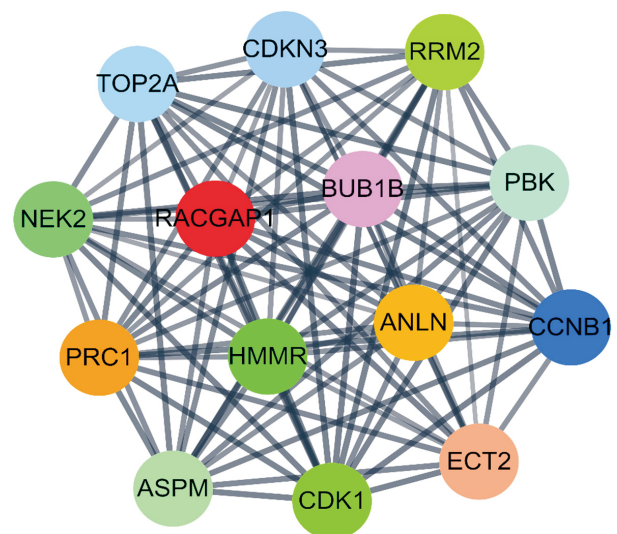


Fig. 7. Protein-protein interaction network of the top 14 hub genes with a high degree of connectivity.

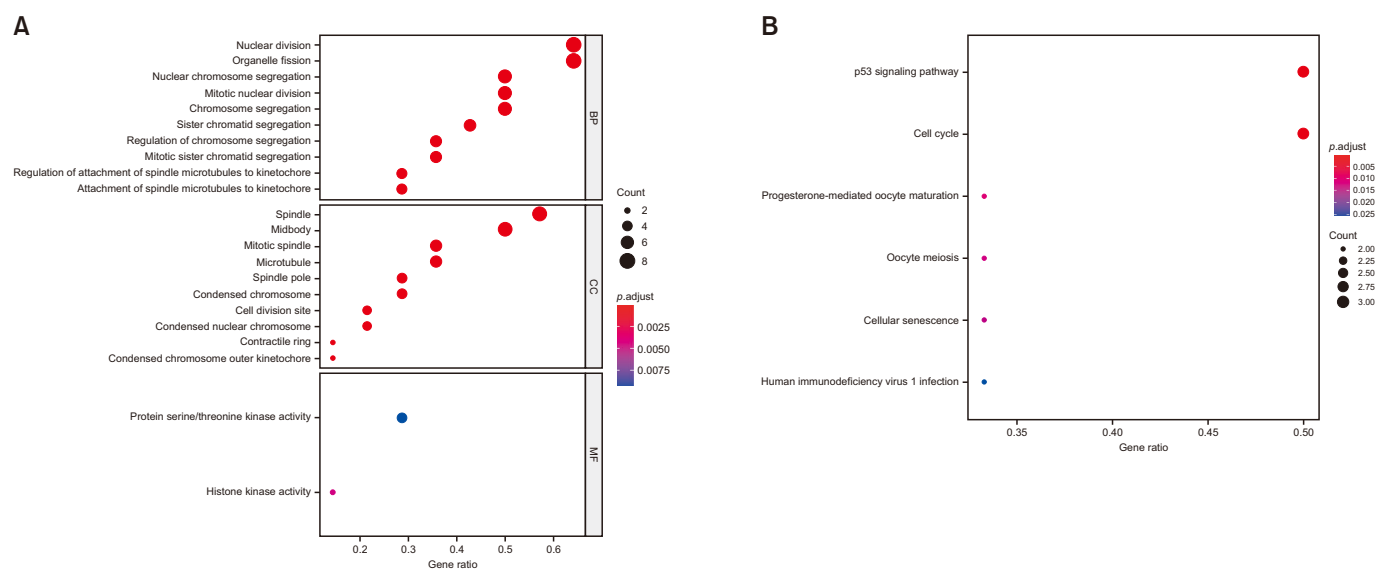


Fig. 8. Enrichment analyses of the 14 hub genes in the most significantly enriched module. Results of gene ontology (A) and Kyoto Encyclopedia of Genes and Genome (B) enrichment analyses are shown.

Table 1. Pathway enrichment analysis of the 14 hub genes in the most significant modules

Category	GO ID	Term	p-value	Count
BP	GO:0000280	Nuclear division	2.65E-12	9
BP	GO:0048285	Organelle fission	6.88E-12	9
BP	GO:0051988	Regulation of attachment of spindle microtubules to kinetochores	1.35E-10	4
BP	GO:0098813	Nuclear chromosome segregation	3.77E-10	7
BP	GO:0140014	Mitotic nuclear division	5.36E-10	7
BP	GO:0007059	Chromosome segregation	1.58E-09	7
BP	GO:0000819	Sister chromatid segregation	3.26E-09	6
BP	GO:0008608	Attachment of spindle microtubules to kinetochores	8.69E-09	4
BP	GO:0051983	Regulation of chromosome segregation	1.03E-08	5
BP	GO:0000070	Mitotic sister chromatid segregation	8.00E-08	5
CC	GO:0030496	Midbody formation	1.75E-11	7
CC	GO:0005819	Spindle formation	3.88E-11	8
CC	GO:0072686	Mitotic spindle formation	2.57E-08	5
CC	GO:0000922	Spindle pole formation	4.58E-06	4
CC	GO:0005874	Microtubule formation	8.62E-06	5
CC	GO:0000793	Chromosome condensation	1.48E-05	4
CC	GO:0032153	Cell division site	1.69E-05	3
CC	GO:0070938	Contractile ring formation	2.13E-05	2
CC	GO:0000940	Chromosome condensation on outer kinetochores	4.31E-05	2
CC	GO:0000794	Nuclear chromosome condensation	4.95E-05	3
MF	GO:0035173	Histone kinase activity	7.30E-05	2
MF	GO:0004674	Protein serine/threonine kinase activity	0.000258	4
KEGG	hsa04115	p53 signaling pathway	1.38E-05	3
KEGG	hsa04110	Cell cycle	6.80E-05	3
KEGG	hsa04914	Progesterone-mediated oocyte maturation	0.002199	2
KEGG	hsa04114	Oocyte meiosis	0.003632	2
KEGG	hsa04218	Cellular senescence	0.005271	2
KEGG	hsa05170	Human immunodeficiency virus 1 infection	0.009571	2

GO, gene ontology; BP, biological process; CC, cellular component; MF, molecular function; KEGG, Kyoto Encyclopedia of Genes and Genome.

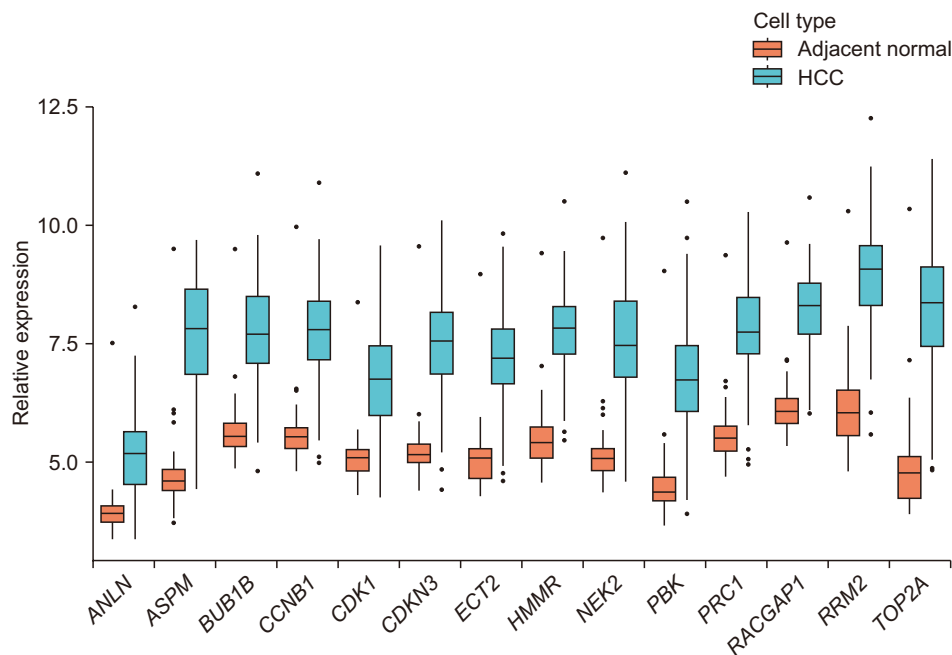


Fig. 9. Statistical analysis of expression of the 14 hub genes in GSE121248. HCC, hepatocellular carcinoma.

close associations between expression levels of the 14 hub genes and the overall survival of HCC patients. High expression levels of 12 hub genes (all 14 hub genes except *CDKN3* and *RACGAP1*) were strongly correlated with poor patient prognosis ($p < 0.05$).

DISCUSSION

HCC is currently the most common type of primary liver cancer in adults. It is also the most frequent cause of death in patients with liver cirrhosis. HCC in many countries, including Korea, is closely related to HBV infection. Although patients infected with HBV are regularly screened for HCC by measuring blood levels of alpha-fetoprotein (AFP) and abdominal ultrasonography every six months, HBV-associated HCC remains the main cause of mortality in HBV-infected patients [4]. The identification of new biomarkers and regulatory pathways related to the development of HCC in HBV-infected patients can enhance early detection, diagnosis, and treatment of HCC. The present study utilized bioinformatics to analyze the GSE121248 gene expression dataset and screen for key genes and molecular pathways that are closely related to the development of HCC. Using cut-off criteria of $p < 0.05$ and $|\log_2 FC| > 2$, this analysis identified 134 DEGs, of which 100 were down-regulated and 34 were up-regulated in HCC than in normal liver tissues. To obtain a comprehensive understanding of these DEGs, we performed bioinformatics analyses, including GO enrichment, KEGG pathway, and PPI network analyses. To validate the 14 hub genes identified, we also analyzed their expression levels using the GEPIA website and assessed their effects on overall patient survival using the TCGA database.

GO enrichment analysis revealed that the 34 up-regulated

DEGs were significantly enriched in genes involved in nuclear division and organelle fission, whereas the 100 down-regulated DEGs were highly enriched in genes involved in steroid metabolism. KEGG pathway analysis showed that the 34 up-regulated genes were enriched in genes involved in the p53 signaling pathway, whereas the 100 down-regulated genes were significantly enriched in genes involved in retinol metabolism. PPI network analysis was performed using STRING. Hub genes were identified by MCODE, a Cytoscape plug-in. The 14 hub genes identified by PPI network analysis included *ANLN*, *ASPM*, *BUB1B*, *CCNB1*, *CDK1*, *CDKN3*, *ECT2*, *HMMR*, *NEK2*, *PBK*, *PRC1*, *RACGAP1*, *RRM2*, and *TOP2A*. Analysis using the GEPIA website showed that expression levels of all 14 mRNAs were much higher in HCC than in normal tissue samples. Survival analysis using the TCGA database showed that expression levels of these genes were significantly associated with patient survival.

Anillin (ANLN) is a conserved protein implicated in cytoskeletal dynamics during cytokinesis and cellularization. ANLN is essential for cell division. Therefore, it is critical for development and homeostasis in metazoans. ANLN expression levels are correlated with the metastatic potential of human tumors. For example, colorectal cancer cell lines expressing higher levels of ANLN replicated faster, were more invasive, and had greater levels of migration than cell lines expressing lower levels of ANLN [5]. Overexpression of ANLN may enhance cell migration by cytoskeletal remodeling, promoting the proliferation, invasion, and motility of tumor cells [6]. In addition, overexpression of ANLN involving copy number variation, as well as alterations in DNA and histone methylation, was found to be related to poor prognosis in patients with HCC [7].

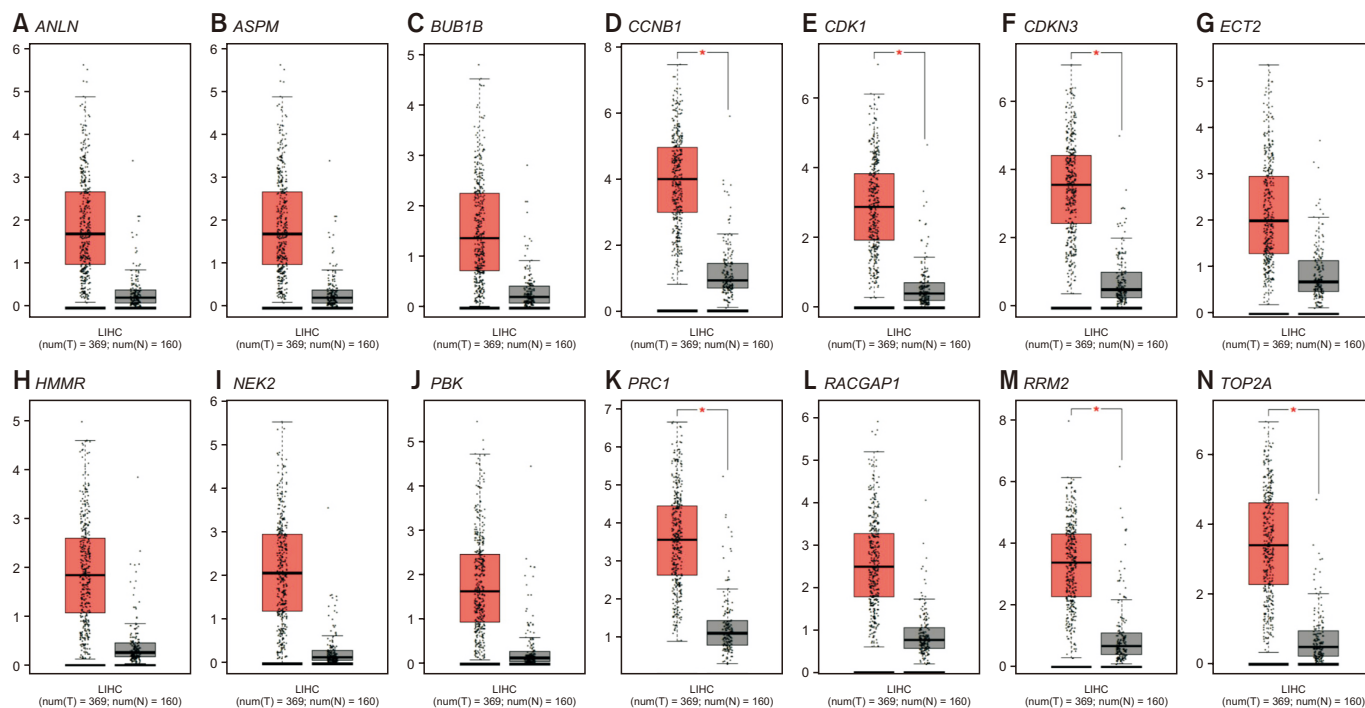


Fig. 10. Expression levels of mRNAs encoded by the 14 hub genes based on TCGA database. (A–N) *ANLN*, *ASPM*, *BUB1B*, *CCNB1*, *CDK1*, *CDKN3*, *ECT2*, *HMMR*, *NEK2*, *PBK*, *PRC1*, *RACGAP1*, *RRM2*, and *TOP2A*, respectively. LIHC, liver hepatocellular carcinoma; T, tumor; N, normal. * p -value < 0.05 versus normal.

The abnormal spindle-like microcephaly-associated (*ASPM*) protein plays a significant role in neurogenesis and cell proliferation. In addition, *ASPM* overexpression has been associated with poor outcomes of patients with various cancers [8,9]. *ASPM* overexpression may predict HCC invasiveness or metastatic potential. It is associated with a higher risk of early tumor recurrence regardless of p53 mutation status or tumor stage [10].

The mitotic checkpoint serine/threonine-protein kinase *BUB1* beta (*BUB1B*) gene encodes a kinase involved in spindle checkpoint function and chromosome segregation. Impaired spindle checkpoint function is involved in the initiation and development of tumors. In addition, *BUB1B* can promote HCC progression by activating the mTORC1 signaling pathway [11].

Cyclin B1 (*CCNB1*) is a regulatory protein involved in mitosis. A high level of *CCNB1* expression may be associated with the extent of tumor progression and invasion, suggesting that the expression level of *CCNB1* might be prognostic of survival outcomes in cancer patients. Transcriptional activation of *CCNB1* is essential for the proliferation of HCC. It is closely associated with poor prognosis in patients with HCC [12].

Cyclin-dependent kinase 1 (*CDK1*), a member of the Ser/Thr protein kinase family, plays an important role in G1/S and G2/M phase transitions of eukaryotic cell cycle by interacting with *CCNB1*. *CDK1* was found to play an essential role in HCC development by regulating subcellular localization of apoptin

[13]. Metformin can significantly inhibit HCC cell proliferation by inducing G2/M arrest and reducing the expression of *CDK1*, suggesting that *CDK1* might be involved in the process of HCC cell proliferation [14]. In addition, overexpression of *CDK1* in HCC is significantly associated with large tumor size, advanced tumor stage, portal invasion, intrahepatic metastasis, poor differentiation, high AFP level, high Ki-67 index, and poor patient prognosis [15].

Cyclin-dependent kinase inhibitor 3 (*CDKN3*), which inhibits the activation of cyclin-dependent kinase 2 (*CDK2*), can promote tumorigenesis in various types of cancer [16]. In contrast to our findings, a recent study has found that *CDKN3* is negatively associated with pathological tumor stage and that its inhibition can enhance HCC cell survival [17].

The epithelial cell transformation sequence 2 (*ECT2*) protein is a Rho-specific exchange factor. Activation of the *ECT2*/Rho pathway can enhance the progression of several tumors. *ECT2* overexpression may promote the progression and early recurrence of HCC by regulating the Rho/ERK signaling axis [18]. High levels of *ECT2* have been implicated in other types of cancer [19].

Hyaluronan-mediated motility receptor (*HMMR* or *RHAMM*) was originally identified as a soluble protein that altered migratory cell behavior and bound to hyaluronan. *HMMR* expression is associated with breast cancer risk and the progression of several tumor types. A recent bioinformatics

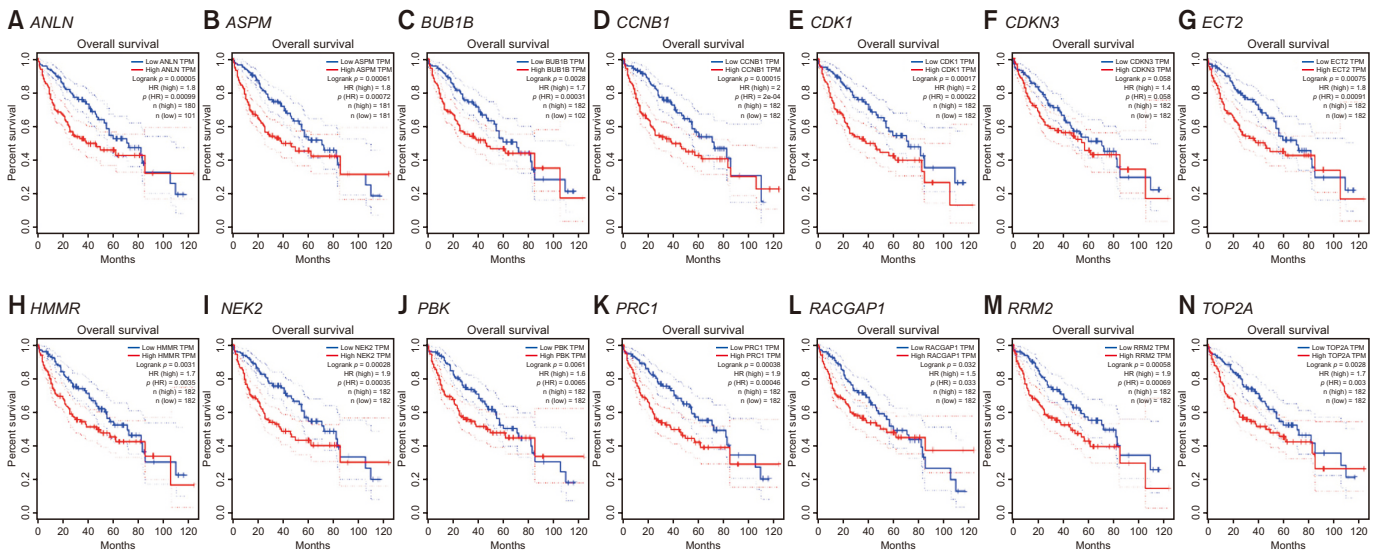


Fig. 11. Kaplan-Meier analysis of overall patient survival relative to the expression of the 14 hub genes based on the TCGA database. (A–N) *ANLN*, *ASPM*, *BUB1B*, *CCNB1*, *CDK1*, *CDKN3*, *ECT2*, *HMMR*, *NEK2*, *PBK*, *PRC1*, *RACGAP1*, *RRM2*, and *TOP2A*, respectively. T, tumor; N, normal. * p -value < 0.05 versus normal.

study has found that overexpression of HMMR is positively correlated with HCC tumor grade, pathological stage, and tumor stage [20].

The serine/threonine-protein kinase Nek2 (NEK2) is a microtubule-associated protein that regulates spindle assembly in human cells. It is overexpressed in malignancies. NEK2 up-regulation has been observed in primary HCC tissue and HCC cell lines. It is significantly associated with cancer progression, drug resistance, and poor patient prognosis [21]. In contrast to these findings, another study has reported that low expression of NEK2 is associated with HCC progression and poor patient prognosis [22].

Lymphokine-activated killer T-cell-originated protein kinase (PBK) is a serine/threonine-protein kinase associated with the dual specific mitogen-activated protein kinase kinase (MAPKK) family. PBK is involved in the activation of lymphoid cells, which support testicular function and spermatogenesis. Overexpression of PBK has been associated with tumor progression. In HCC, PBK overexpression can promote metastasis through activation of the ETV4-uPAR signaling pathway [23].

Protein regulator of cytokinesis 1 (PRC1) is a protein that regulates antiparallel microtubule cross-linking, promoting the formation of microtubules that support cell shape and regulate cytokinesis. PRC1 is overexpressed in HCC. Its high expression correlates with early HCC recurrence and chemoresistance [24].

Rac GTPase-activating protein 1 (RACGAP1) belongs to the family of Rho GTPase-activating proteins. It plays a key role in controlling the initiation of cytokinesis, transformation, invasive migration, and metastasis. Overexpression of RacGAP1 protein in colorectal cancer is significantly associated with higher T stage, vessel invasion, and distant metastasis. It is

therefore significantly related to poor disease-free and overall survival rates [25]. In HCC, overexpression of RACGAP1 correlates with early recurrence and a high migration rate [26].

Ribonucleotide reductase subunit M2 (RRM2) catalyzes the production of deoxynucleotide triphosphates, which are essential for both replicative and repair DNA synthesis. RRM2 is known to play an active role in tumor progression. Elevated RRM2 levels are correlated with poorer prognosis in patients with several types of cancer [27]. Overexpression of RRM2 in HCC correlates with early recurrence and intrahepatic metastasis. It is therefore associated with poor prognosis [28].

The DNA topoisomerase 2 α (*TOP2A*) gene located on human chromosome 17 encodes the enzyme DNA topoisomerase II α . *TOP2A* expression may be a significant predictor of tumor progression, recurrence, and poor survival rate in various types of cancer [29]. In HCC, high expression of *TOP2A* has been associated with more advanced clinical stages, lower grades of tumor differentiation, high T stages, and resistance to chemotherapy [30].

Patterns of expression of the 14 hub genes identified in this study were similar in other malignancies. In addition, several *in vivo*, *in vitro*, and bioinformatics studies have shown that individual expression level of each of these 14 hub genes is correlated with the development and progression of HCC.

The background purpose of the present study was to identify key genes and pathways of HBV-associated HCC as a preliminary study before actual experimental study using tissue samples obtained from our own patients. It is essential to perform external validation studies focusing on results of the present study.

The present study has limitations. This study included only

one public database. It is necessary to validate results of this study with external data sets obtained from patients of our institution and other Korean institutions.

In conclusion, this comprehensive bioinformatics analysis identified candidate genes and pathways that might be involved in the progression of HBV-associated HCC. These findings may provide better understanding of the molecular mechanisms and potential biomarkers of HCC. Although further experiments with additional patients are required to confirm these findings, results of the present study may suggest future genomic treatment targets for patients with HCC.

FUNDING

None.

CONFLICT OF INTEREST

No potential conflict of interest relevant to this article was reported.

ORCID

Sang-Hoon Kim, <https://orcid.org/0000-0002-8025-1816>
 Shin Hwang, <https://orcid.org/0000-0002-9045-2531>
 Gi-Won Song, <https://orcid.org/0000-0002-4235-0434>
 Dong-Hwan Jung, <https://orcid.org/0000-0001-5984-023X>
 Deok-Bog Moon, <https://orcid.org/0000-0002-8209-3540>
 Jae Do Yang, <https://orcid.org/0000-0001-9701-7666>
 Hee Chul Yu, <https://orcid.org/0000-0003-2766-1354>

AUTHOR CONTRIBUTIONS

Conceptualization: SH, SHK. Data curation: SHK, GWS. Methodology: DHJ, DBM, JDY, HCY. Visualization: SHK, SH. Writing - original draft: SHK, SH. Writing - review & editing: All authors.

REFERENCES

- Singal AG, Lampertico P, Nahon P. Epidemiology and surveillance for hepatocellular carcinoma: new trends. *J Hepatol* 2020;72:250-261.
- Liu S, Yao X, Zhang D, Sheng J, Wen X, Wang Q, et al. Analysis of transcription factor-related regulatory networks based on bioinformatics analysis and validation in hepatocellular carcinoma. *Biomed Res Int* 2018;2018:1431396.
- Shen S, Kong J, Qiu Y, Yang X, Wang W, Yan L. Identification of core genes and outcomes in hepatocellular carcinoma by bioinformatics analysis. *J Cell Biochem* 2019;120:10069-10081.
- Forner A, Llovet JM, Bruix J. Hepatocellular carcinoma. *Lancet* 2012;379:1245-1255.
- Zhang L, Maddox AS. Anillin. *Curr Biol* 2010;20:R135-R136.
- Chuang HY, Ou YH. Abstract 4068: overexpression of anillin in colorectal cancer promotes the cell proliferation, cell mobility and cell invasion. *Cancer Res* 2014;74(19 Suppl):4068.
- Zhang LH, Wang D, Li Z, Wang G, Chen DB, Cheng Q, et al. Overexpression of anillin is related to poor prognosis in patients with hepatocellular carcinoma. *Hepatobiliary Pancreat Dis Int* 2021;20:337-344.
- Xu Z, Zhang Q, Luh F, Jin B, Liu X. Overexpression of the ASPM gene is associated with aggressiveness and poor outcome in bladder cancer. *Oncol Lett* 2019;17:1865-1876.
- Xie JJ, Zhuo YJ, Zheng Y, Mo RJ, Liu ZZ, Li BW, et al. High expression of ASPM correlates with tumor progression and predicts poor outcome in patients with prostate cancer. *Int Urol Nephrol* 2017;49:817-823.
- Lin SY, Pan HW, Liu SH, Jeng YM, Hu FC, Peng SY, et al. ASPM is a novel marker for vascular invasion, early recurrence, and poor prognosis of hepatocellular carcinoma. *Clin Cancer Res* 2008;14:4814-4820.
- Qiu J, Zhang S, Wang P, Wang H, Sha B, Peng H, et al. BUB1B promotes hepatocellular carcinoma progression via activation of the mTORC1 signaling pathway. *Cancer Med* 2020;9:8159-8172.
- Chai N, Xie HH, Yin JP, Sa KD, Guo Y, Wang M, et al. FOXM1 promotes proliferation in human hepatocellular carcinoma cells by transcriptional activation of CCNB1. *Biochem Biophys Res Commun* 2018;500:924-929.
- Zhao J, Han SX, Ma JL, Ying X, Liu P, Li J, et al. The role of CDK1 in apoptin-induced apoptosis in hepatocellular carcinoma cells. *Oncol Rep* 2013;30:253-259.
- Zhou J, Han S, Qian W, Gu Y, Li X, Yang K. Metformin induces miR-378 to downregulate the CDK1, leading to suppression of cell proliferation in hepatocellular carcinoma. *Onco Targets Ther* 2018;11:4451-4459.
- Ito Y, Takeda T, Sakon M, Monden M, Tsujimoto M, Matsuura N. Expression and prognostic role of cyclin-dependent kinase 1 (cdc2) in hepatocellular carcinoma. *Oncology* 2000;59:68-74.
- Fan C, Chen L, Huang Q, Shen T, Welsh EA, Teer JK, et al. Overexpression of major CDKN3 transcripts is associated with poor survival in lung adenocarcinoma. *Br J Cancer* 2015;113:1735-1743.
- Dai W, Miao H, Fang S, Fang T, Chen N, Li M. CDKN3 expression is negatively associated with pathological tumor stage and CDKN3 inhibition promotes cell survival in hepatocellular carcinoma. *Mol Med Rep* 2016;14:1509-1514.
- Chen J, Xia H, Zhang X, Karthik S, Pratap SV, Ooi LL, et al. ECT2 regulates the Rho/ERK signalling axis to promote early recurrence in human hepatocellular carcinoma. *J Hepatol* 2015;62:1287-1295.
- Jin Y, Yu Y, Shao Q, Ma Y, Zhang R, Yao H, et al. Up-regulation of ECT2 is associated with poor prognosis in gastric cancer patients. *Int J Clin Exp Pathol* 2014;7:8724-8731.
- He Z, Mei L, Connell M, Maxwell CA. Hyaluronan mediated motility receptor (HMMR) encodes an evolutionarily conserved homeostasis, mitosis, and meiosis regulator rather than a hyaluronan receptor. *Cells* 2020;9:819.
- Lin S, Zhou S, Jiang S, Liu X, Wang Y, Zheng X, et al. NEK2 regulates stem-like properties and predicts poor prognosis in hepatocellular

- carcinoma. *Oncol Rep* 2016;36:853-862.
22. Fu L, Liu S, Wang H, Ma Y, Li L, He X, et al. Low expression of NEK2 is associated with hepatocellular carcinoma progression and poor prognosis. *Cancer Biomark* 2017;20:101-106.
 23. Yang QX, Zhong S, He L, Jia XJ, Tang H, Cheng ST, et al. PBK overexpression promotes metastasis of hepatocellular carcinoma via activating ETV4-uPAR signaling pathway. *Cancer Lett* 2019;452:90-102.
 24. Wang Y, Shi F, Xing GH, Xie P, Zhao N, Yin YF, et al. Protein regulator of cytokinesis PRC1 confers chemoresistance and predicts an unfavorable postoperative survival of hepatocellular carcinoma patients. *J Cancer* 2017;8:801-808.
 25. Imaoka H, Toiyama Y, Saigusa S, Kawamura M, Kawamoto A, Okugawa Y, et al. RacGAP1 expression, increasing tumor malignant potential, as a predictive biomarker for lymph node metastasis and poor prognosis in colorectal cancer. *Carcinogenesis* 2015;36:346-354.
 26. Wang SM, Ooi LL, Hui KM. Upregulation of Rac GTPase-activating protein 1 is significantly associated with the early recurrence of human hepatocellular carcinoma. *Clin Cancer Res* 2011;17:6040-6051.
 27. Nana AW, Wu SY, Yang YS, Chin YT, Cheng TM, Ho Y, et al. Nano-diamino-tetrac (NDAT) enhances resveratrol-induced antiproliferation by action on the RRM2 pathway in colorectal cancers. *Horm Cancer* 2018;9:349-360.
 28. Lee B, Ha SY, Song DH, Lee HW, Cho SY, Park CK. High expression of ribonucleotide reductase subunit M2 correlates with poor prognosis of hepatocellular carcinoma. *Gut Liver* 2014;8:662-668.
 29. Zhao Q, Li H, Zhu L, Hu S, Xi X, Liu Y, et al. Bioinformatics analysis shows that TOP2A functions as a key candidate gene in the progression of cervical cancer. *Biomed Rep* 2020;13:21.
 30. Wong N, Yeo W, Wong WL, Wong NL, Chan KY, Mo FK, et al. TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance. *Int J Cancer* 2009;124:644-652.