Conjoint analysis for hepatic carcinoma with hub genes and multi-slice spiral CT

Shuang Zhang, MSa, Ruchen Peng, MSa+b, Ruiqiang Xin, MSb, Xiuzhi Shen, MSa, Jingli Zheng, MSa

Abstract
Hepatic carcinoma (HCC) is a common malignant tumor, with insidious onset and poor prognosis. However, more hub genes associated with hepatocellular carcinoma are unknown. And there are few researches about the conjoint analysis with the hub genes and multi-slice spiral computerized tomography (CT).

A total of 100 HCC participates were recruited, who all received the examination of multi-slice spiral CT. Two expression profile data sets (GSE101728 and GSE101685) were downloaded from the Gene Expression Omnibus (GEO) database. GEO2R can perform a command to compare gene expression profiles between groups in order to identify differently expressed genes (DEGs). Functional annotation of DEGs via Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was made with Database for Annotation, Visualization, and Integrated Discovery (DAVID). Construction and analysis of protein–protein interaction network were performed. Furthermore, the study could mine of hub genes and explore the correlation with the multi-slice CT. Real-time quantitative polymerase chain reaction (RT-qPCR) assay was used the exam the expression of hub genes.

A total of 10 genes were identified as hub genes with degrees ≥10. The hub genes (NIMA Related Kinase 2 [NEK2], Anillin Actin Binding Protein [ANLN], DNA Topoisomerase II Alpha [TOP2A], Centromere Protein F [CENPF], Assembly Factor For Spindle Microtubules [ASPM], Cell Division Cycle 20 [CDC20], Cyclin Dependent Kinase 1 [CDK1], Cyclin B1 [CCNB1], Epithelial Cell Transforming 2 [ECT2], Cyclin B2 [CCNB2]) were identified from the Molecular Complex Detection (MCODE) network. These hub genes were highly expressed in HCC tissues, and when these genes were highly expressed, the survival prognosis of HCC patients was poor. The type of CT enhancement was significantly related with the expression of NEK2 (P < .001), ANLN (P < .001), and TOP2A (P = .006).

The combination between the gene expression (NEK2, ANLN, and TOP2A) and type of CT enhancement might provide a new idea for future basic research and targeted therapy of HCC.

Abbreviations: ANLN = Anillin Actin Binding Protein, ASPM = Assembly Factor For Spindle Microtubules, BP = biological processes, CC = cell component, CCNB1 = Cyclin B1, CCNB2 = Cyclin B2, CDC6 = Cell Division Cycle 6, CDC20 = cell division cycle 20, CDK = cyclin dependent kinase, CDK1 = cyclin dependent kinase 1, cDNA = complementary deoxyribonucleic acid, CENP = centromere protein, CENPF = centromere protein F, CT = computerized tomography, DAVID = Database for Annotation, Visualization and Integrated Discovery, DEGs = differently expressed genes, ECT2 = epithelial cell transforming 2, FO = fold change, FoxM1 = Forkhead Box M1, GAPDH = glyceraldehyde-3-phosphate dehydrogenase, GEO = gene expression omnibus, GO = gene ontology, HCC = hepatic carcinoma, KEGG = Kyoto Encyclopedia of Genes and Genomes, KIF20A = Kinesin Family Member 20A, MAPK = mitogen-activated protein kinase, MCODE = Molecular Complex Detection, MF = molecular function, MMP = matrix metallopeptidase, NEK2 = NIMA-related kinase 2, PPI = protein–protein interaction, RNA = ribose nucleic acid, RTKN = Rhotekin, RT-qPCR = real-time quantitative polymerase chain reaction, SOX2 = SRY-Box Transcription Factor 2, STRING = Search Tool for the Retrieval of Interacting Genes, TOP2A = DNA Topoisomerase II Alpha.

Keywords: hepatic carcinoma, multi-slice spiral CT, differentially expressed genes, protein–protein interaction, bioinformatic analysis
1. Introduction

Primary HCC is a common malignant tumor originating from hepatic stroma or epithelial tissue, with insidious onset and poor prognosis. Its high incidence is related to the degree of malignancy. HCC is the sixth most common cancer in the world and the second leading cause of cancer death worldwide. It is a major international health problem.

The American Cancer Society estimates that 42,220 new cases of HCC will be diagnosed in 2018, and 30,200 of those will die. Patients with liver cancer usually do not have typical liver symptoms in the early stage, such as jaundice, liver failure, and ascites, until very late. Risk factors for HCC include hepatitis B or hepatitis C virus infection, alcoholism, and fatty liver disease. The molecular mechanism of primary liver cancer progression has not been fully defined, but it is believed to be closely related to the molecular pathway expression of Matrix Metalloproteinase (MMP)-2, and MMP-9 in cancer tissues and adjacent tissues.

As for the targeted therapy of HCC, sorafenib was used clinically in 2007 as the first molecular-targeted agent. Current studies demonstrate that Regorafenib, Cabozantinib, and Ramucirumab are appropriate supplements for sorafenib as second-line therapy in patients with advanced HCC who are resistant, progressive, or intolerant to sorafenib. Also the increase of SRY-Box Transcription Factor 2 (SOX2) gene expression is related to the occurrence and development of HCC. However, more core genes associated with hepatocellular carcinoma are unknown.

Bioinformatics is a science that studies the processing of biological information (collection, management, and application of analysis) and extracts new knowledge of biology from it. It links biodata with medical science research, which is becoming more and more important in gene discovery of targeted therapies. Tsimeridou found that using bioinformatics analysis can apply new discoveries in cancer biology to the clinic, promising to greatly improve clinical outcomes for cancer patients. Shing-ng et al have demonstrated that understanding the functional impact of somatic mutations in cancer through bioinformatics analysis can improve patient care and treatment more precisely.

Nowadays, imaging diagnosis technology has received much attention in the screening and evaluation of liver cancer. Multi-slice spiral CT is adopted for the examination of liver cancer, which can observe the abnormal image characteristics of the incidence area and realize the differentiation of benign and malignant tumors. To further clarify the characteristics and development of the HCC, Multi-slice spiral CT examination can be combined with gene expression to analyze the characteristics of HCC.

Therefore, this study intends to explore and verify the core genes of HCC using bioinformatics techniques and clinical data, and to research the correlation with multi-slice CT. This study will provide a new direction for exploring the pathogenesis and specific targeted therapy of HCC.

2. Materials and methods

2.1. Patients and ethics

A total of 100 HCC participates (43–75, average age is 56.2±3.12 years old) were recruited, including 55 female patients and 45 male patients. The research conformed to the Declaration of Helsinki and was authorized by the Human Ethics and Research Ethics Committees of Beijing Luhe Hospital. The written informed consents were obtained from all participates.

2.2. The multi-slice CT

Siemens multi-slice spiral CT was selected as diagnostic instrument. Do a good job of pre-examination guidance to ensure that patients understand the CT examination matters needing attention. After the regular scan was completed, the enhanced scan was performed. A 80mL of contrast agent was injected through cubital vein, and the flow rate was 3.0mL/s. All patients underwent phase III scanning. Finally, according to the CT value of tumor parenchyma and tumor-related vessels, tumor parenchyma enhancement, tumor vascular enhancement, and tumor non-enhancement were classified.

2.3. Access to public data

GEO (http://www.ncbi.nlm.nih.gov/geo) is an open high-throughput genomic database that includes microarrays, gene expression data, and chips. Two expression profile data sets (GSE101728 and GSE101685) were downloaded from the GEO database. The GSE101728 data set is composed of 7 HCC tissues and 7 adjacent normal liver tissues. The GSE101685 date set is composed of 24 HCC tissues and 8 adjacent normal liver tissues.

2.4. Screening of DEGs via GEO2R

GEO2R (http://www.ncbi.nlm.nih.gov/geo/geo2r) is a system for online analysis of data in GEO, which runs in the R language. GEO2R can perform a command to compare gene expression profiles between groups in order to identify DEGs between HCC and normal groups. In general, when the probe set has a corresponding gene symbol, the probe is considered valuable and will be retained. Statistically significant measure is P-value ≤0.01 and Fold change (FC) ≥3.

2.5. Functional annotation of DEGs via GO and KEGG analysis

DAVID (https://david.ncifcrf.gov/home.jsp) (version 6.8) is a bioinformatics database that integrates biological data and analytical tools to provide systematic and comprehensive annotation information of biological functions for large-scale genes or protein lists to help users extract biological information from them. KEGG (https://www.kegg.jp/) is one of the most commonly used bioinformatics databases in the world to understand advanced functions and biological systems. At the molecular level, KEGG integrates a large number of utility database resources from high-throughput experimental techniques. GO is an ontology widely used in bioinformatics, which covers three aspects of biology, including cellular components, molecular function, and biological process. In order to analyze the GO and pathway enrichment information of DEGs, the DAVID online tool was executed. Statistically significant measure is P < 0.05.

Furthermore, the Metascape was performed to make functional annotation to verify the results of DAVID.

2.6. Construction and analysis of protein–protein interaction network

Search Tool for the Retrieval of Interacting Genes (STRING; http://string-db.org) (version 10.5) is a network that can be used to predict and track protein–protein interactions (PPIs). In this study, STRING database was used to construct PPI network with
DEGs. Cytoscape (version 3.6.1) is open visualization software that can be used to visualize PPI network. Based on topological principles, the MCODE (version 1.5.1), a plug-in for Cytoscape, can mine tightly coupled regions from PPI. First, Cytoscape software plots the PPI network. Secondly, MCODE identifies the most important modules in the PPI network graph. The criteria of MCODE analysis is that node score cut-off = 0.2, degree cut-off = 2, Max depth = 100, MCODE scores >3, and k-score = 2.

### 2.7. Mining and screening of hub gene

The hub genes were selected with degrees ≥10. Hierarchical clustering of hub genes was constructed using the expression level of them in the datasets through the R language. The overall survival of hub genes was performed using Kaplan–Meier curve in Kaplan–Meier Plotter (http://kmplot.com/analysis/index.php?p=background).

### 2.8. RT-qPCR assay

Total Ribose Nucleic Acid (RNA) was extracted by the RNAiso Plus (Trizol) kit (Thermofisher, MA), and reverse transcribed to complementary deoxyribonucleic acid (cDNA). RT-qPCR was performed using a Light Cycler 4800 System with specific primers for genes. The relative quantitation values (\(2^{-\Delta\Delta C_T}\), where \(C_T\) is the threshold cycle) of each sample were calculated, and are presented as fold change in gene expression relative to the control group. Glyceraldehyde-3-Phosphate Dehydrogenase (GAPDH) was used as an endogenous control. A total of 10 HCC tumor samples from HCC patients and 10 adjacent normal liver tissues samples were selected randomly to make comparison about the expression of hub genes.

### 2.9. Statistical analysis

Student’s t test was used to determine the statistical significance when comparing the two groups. Enumeration data were expressed as rate, and Pearson chi-square test was used to explore the relationship between the multi-slice CT and the expression of hub genes.

### 3. Results

#### 3.1. Identification of DEGs in HCC

One volcano plot presents the DEGs in the GSE101728 (Fig. 1A), and another volcano plot presents the DEGs in the GSE101685 (Fig. 1B). After standardization of the microarray results, DEGs were identified. The overlap among the 2 datasets contained 79 genes as shown in the Venn diagram (Fig. 1C).

#### 3.2. KEGG and GO enrichment analyses of DEGs

GO analysis results showed that changes in biological processes (BP) of DEGs were significantly enriched in positive regulation of cell proliferation, negative regulation of ubiquitin–protein ligase activity involved in mitotic cell cycle, mitotic spindle assembly checkpoint, cell division and so on (Fig. 2A). Changes in cell component (CC) of DEGs were mainly enriched in organelle membrane, midbody, extracellular region, and so on (Fig. 2B). Changes in molecular function (MF) were mainly enriched in arachidonic acid epoxygenase activity, oxido-reductase activity, steroid hydroxylase activity, and so on (Fig. 2C). KEGG pathway analysis showed that all DEGs are mainly concentrated in Drug metabolism-cytochrome P450, Chemical carcinogenesis, Retinol metabolism, and so on (Fig. 2D).

The results of Metascape also demonstrated that the DEGs were mainly enriched in the relative terms of HCC (Fig. 3).

#### 3.3. Protein–protein interaction network construction and module analysis

The PPI network of DEGs was constructed (Fig. 4A) and the most significant module was obtained using Cytoscape (Fig. 4B).

#### 3.4. Hub gene selection and analysis

A total of 10 genes were identified as hub genes with degrees ≥10. The hub genes (NEK2, ANLN, TOP2A, CENPF, ASPM, CDC20, CDK1, CCNB1, ECT2, and CCNB2) were identified from the MCODE network (Fig. 4C). Hierarchical clustering showed that the hub genes could basically differentiate the HCC samples from the non-cancerous samples (Fig. 4D and E). Subsequently, the overall survival analysis of the hub genes was performed using Kaplan–Meier curve. HCC patients with high expression of these hub genes showed worse overall survival (Fig. 5).

#### 3.5. Results of RT-qPCR analysis

According to the above expression analysis, these hub genes were markedly up-regulated in HCC tumor samples. As presented in Figure 6, the relative expression levels of hub genes were significantly higher in HCC samples, compared with the normal liver tissues groups. The result demonstrated that these hub genes might be considered as biomarkers for HCC.

#### 3.6. The correlation between the gene expression and type of multi-slice CT enhancement

Based on the Pearson’s chi-squared test, type of CT enhancement was significantly related with the expression of NEK2 (\(P < .001\), ANLN (\(P < .001\), and TOP2A (\(P = .006\). When the NEK2, ANLN, and TOP2A were highly expressed in the HCC, the percentage of tumor parenchymal was high (Table 1).

### 4. Discussion

In this study, bioinformatics techniques and clinical data were used to verify that NEK2, ANLN, TOP2A, CENPF, ASPM, CDC20, CDK1, CCNB1, ECT2, and CCNB2 were highly expressed in HCC tissues, and when these genes were highly expressed, the survival prognosis of liver cancer patients was poor. The type of CT enhancement was significantly related with the expression of NEK2 (\(P < .001\), ANLN (\(P < .001\), and TOP2A (\(P = .006\).

NEK2 is a protein coding gene. Related pathways include regulation of PLK1 activity during G2/M transition and Cyclin Dependent Kinase (CDK)-mediated phosphorylation and removal of Cell Division Cycle 6 (CDC6). Wonesy et al concluded through microarray analysis that Forkhead Box M1 (FoxM1)
regulates genes that are essential for faithful chromosome separation and mitosis, including NEK2, Kinesin Family Member 20A (KIF20A) and Centromere Protein (CENP)-A. Loss of FoxM1 expression leads to mitotic spindle defects, delayed mitotic cells, and mitotic mutations.\textsuperscript{[14]} Therefore, we inferred that NEK2 could promote the progression of HCC cells by affecting cell proliferation and division. Subsequently, Zhang et al concluded that NEK2 may be a valuable oncogenic factor and a promising target for the treatment of primary liver cancer. NEK2 may regulate the proliferation and apoptosis of HepG2 cells through Mitogen-Activated Protein Kinase (MAPK) signaling pathway.\textsuperscript{[15]} Drozdov et al have shown that the expression of

Figure 1. DEGs in HCC. (A) One volcano plot presents the DEGs in the GSE101728. (B) Another volcano plot presents the DEGs in the GSE101685. (C) Venn diagram.
ASPM, NEK2, and TOP2A genes is closely related to the progression of HCC, and may be used to further identify new targets for effective therapy or diagnosis. Huang et al. investigated that the increased expression of NEK2 is a promising molecular target for new anticancer drugs. ANLN is a protein coding gene. An important by-product of this gene is Rhotekin (RTKN). RTKN, as a cancer promoter, is frequently found in many human cancers, including lung, stomach, colorectal, and bladder cancers. As an actin binding protein, this gene plays a role in cell growth, migration and cytokinesis. Zhou et al. demonstrated that RTKN significantly increased HCC cells and tissues in HCC patients. In addition, it was also found that RTKN is a direct gene target of Mir-152, and mirNA-152 can inhibit the growth of HCC tumor cells by targeting RTKN. TOP2A is a protein coding gene. The related pathways include protein metabolism and proteolysis, which is speculated as sumO-1 pathway. The gene encodes a DNA topoisomerase that controls and changes the topological state of DNA during transcription. Mutations in the gene that encodes this enzyme, which is a target for several anticancer drugs, have been linked to the development of drug resistance. Reduced activity of this
Figure 3. The metascape analysis. (A) Bar graph of enriched terms across input gene lists, colored by $P$-values. (B) Network of enriched terms: colored by cluster ID, where nodes that share the same cluster ID are typically close to each other. (C) Network of enriched terms: colored by $P$-value, where terms containing more genes tend to have a more significant $P$-value. (D) Summary of enrichment analysis in DisGeNET. (E) Summary of enrichment analysis in PaGenBase.
Figure 4. The construction of PPI, MCODE, hub genes networks, and the analysis for the expression of hub genes. (A) The PPI network. (B) The MCODE network. (C) The hub genes network. (D) The heat map for expression of hub genes in the GSE101728. (E) The heat map for expression of hub genes in the GSE101685.
enzyme may also play a role in maladjustment of vasodilatation. Wu et al demonstrated that high expression of TOP2A genes was associated with poor overall survival in HCC patients.\textsuperscript{[16,23]} Ang et al found that TOP2A was over expressed in 25% to 83% liver cancer samples through bioinformatics analysis.\textsuperscript{[24]} The results of Sudan et al showed an Apple flavonoid component as a highly effective TOP2A, which may be a legitimate reason for driving hepatocyte apoptosis.\textsuperscript{[25]}
5. Conclusion

In conclusion, NEK2, ANLN, and TOP2A might play a crucial role in the development of HCC and are expected to become new targets for the treatment of HCC. The combination between the gene expression and type of CT enhancement might provide a new idea for future basic research and targeted therapy.
Acknowledgments

We are thankful to Qiang Dong (Department of Radiology, Shunyi District Hospital) for his assistance in the gene detection.

Author contributions

Conceptualization: Shuang Zhang, Ruiqiang Xin, Xiuzhi Shen. 
Data curation: Ruiqiang Xin, Jingli Zheng. 
Formal analysis: Shuang Zhang, Ruchen Peng. 
Investigation: Ruiqiang Xin, Jingli Zheng. 
Methodology: Ruiqiang Xin, Jingli Zheng. 
Resources: Xiuzhi Shen, Jingli Zheng. 
Software: Xiuzhi Shen. 
Writing – original draft: Shuang Zhang, Ruchen Peng. 
Writing – review & editing: Shuang Zhang, Ruchen Peng, Xiuzhi Shen.

References

[1] Yen C, Sharma R, Rimassa L, et al. Treatment stage migration maximizes survival outcomes in patients with hepatocellular carcinoma treated with Sorafenib: an observational study. Liver Cancer 2017;6:313–24.

[2] Grandhi MS, Kim AK, Ronnekleiv-Kelly SM, et al. Hepatocellular carcinoma: from diagnosis to treatment. Surg Oncol 2016;25:74–85.

[3] El-Serag HB, Rudolph KL. Hepatocellular carcinoma: epidemiology and molecular carcinogenesis. Gastroenterology 2007;132:2537–76.

[4] Massarweh NN, El-Serag HB. Epidemiology of hepatocellular carcinoma and intrahepatic cholangiocarcinoma. Cancer Control 2017;24:1073274817729245.

[5] Meng W, Li X, Bai Z, et al. Silencing alpha-fetoprotein inhibits VEGF and MMP-2/9 production in human hepatocellular carcinoma cell. PLoS One 2014;9:e90660.

[6] Bai ZT, Wu ZR, Xi LL, et al. Inhibition of invasion by N-trans-feruloyloctopamine via AKT, p38MAPK and EMT related signals in hepatocellular carcinoma cells. Bioorg Med Chem Lett 2017;27:989–93.

[7] Liu Z, Lin Y, Zhang J, et al. Molecular targeted and immune checkpoint therapy for advanced hepatocellular carcinoma. J Exp Clin Cancer Res 2019;38:447.

[8] Zhang X, Shao J, Li X, et al. Docetaxel promotes cell apoptosis and decreases SOX2 expression in CD133 expressing hepatocellular carcinoma stem cells by suppressing the H3K/ AKT signaling pathway. Oncol Rep 2019;41:1067–74.

[9] Tsumberidou AM. Targeted therapy in cancer. Cancer Chemother Pharmacol 2015;76:1113–32.

[10] Ng PK, Li J, Jeong KJ, et al. Systematic functional annotation of somatic mutations in cancer. Cancer Cell 2018;33: 450-62.e10.

[11] Balogh J, Victor D3rd, Asham EH, et al. Hepatocellular carcinoma: a review. J Hepatocell Carcinoma 2016;1:41–33.

[12] Long X, Cao J, Shi L, et al. MSCT perfusion imaging and its correlation with perfusion parameters, survivin expression, MVD, and pathologic grade in hepatocellular carcinomas. Zhong Nan Da Xue Xue Bao Yi Xue Ban 2009;34:1096–102.
[13] Ma GL, Jiang HJ, Chen M. Hemodynamic study of hepatocellular carcinoma nodules by multi-slice spiral computed tomographic perfusion. Zhonghua Yi Xue Za Zhi 2013;93:1146–9.

[14] Wonse DR, Follette MT. Loss of the forkhead transcription factor FoxM1 causes centrosome amplification and mitotic catastrophe. Cancer Res 2005;65:5181–9.

[15] Zhang MX, Xu XM, Zhang P, et al. Effect of silencing NEK2 on biological behaviors of HepG2 in human hepatoma cells and MAPK signal pathway. Tumour Biol 2016;37:2023–35.

[16] Drozdov I, Bornschein J, Wex T, et al. Functional and topological properties in hepatocellular carcinoma transcriptome. PLoS One 2012;7:e35510.

[17] Huang LY, Chang CC, Lee YS, et al. Inhibition of Hec1 as a novel approach for treatment of primary liver cancer. Cancer Chemother Pharmacol 2014;74:511–20.

[18] Fan J, Ma LJ, Xia SJ, et al. Association between clinical characteristics and expression abundance of RTKN gene in human bladder carcinoma tissues from Chinese patients. J Cancer Res Clin Oncol 2005;131:157–62.

[19] Qu GQ, Lu YM, Liu YF, et al. Effect of RTKN on progression and metastasis of colon cancer in vitro. Biomed Pharmacother 2015;74:117–23.

[20] Zhang W, Liang Z, Li J. Inhibition of rhotekin exhibits antitumor effects in lung cancer cells. Oncol Rep 2016;35:2529–34.

[21] Sun MY, Zhang H, Tao J, et al. Expression and biological function of rhotekin in gastric cancer through regulating p53 pathway. Cancer Manag Res 2019;11:1069–80.

[22] Zhou J, Zhang Y, Qi Y, et al. MicroRNA-152 inhibits tumor cell growth by directly targeting RTKN in hepatocellular carcinoma. Oncol Rep 2017;37:1227–34.

[23] Wu M, Liu Z, Li X, et al. Analysis of potential key genes in very early hepatocellular carcinoma. World J Surg Oncol 2019;17:77.

[24] Ang C, Miura JT, Gamblin TC, et al. Comprehensive multiplatform biomarker analysis of 350 hepatocellular carcinomas identifies potential novel therapeutic options. J Surg Oncol 2016;113:55–61.

[25] Sudan S, Rupasinghe HP. Flavonoid-enriched apple fraction AF4 induces cell cycle arrest, DNA topoisomerase II inhibition, and apoptosis in human liver cancer HepG2 cells. Nutr Cancer 2014;66:1237–46.