Iran J Public Health, Vol. 53, No.6, Jun 2024, pp.1332-1342

Original Article

A Bioinformatics-Based Approach to Discover Novel Biomarkers in Hepatocellular Carcinoma

Amir Shourideh [1], *Reza Maddah [2], Bahareh Shateri Amiri [3], Zarrin Basharat [4], Marzieh Shadpirouz [5]

1. *Faculty of Pharmacy, Eastern Mediterranean University, Famagusta, Cyprus*
2. *Department of Bioprocess Engineering, Institute of Industrial and Environmental Biotechnology, National Institute of Genetic Engineering and Biotechnology, Tehran, Iran*
3. *Department of Internal Medicine, School of Medicine Hazrat-e Rasool General Hospital Iran University of Medical Sciences, Tehran, Iran*
4. *Alpha Genomics Private Limited, Islamabad 45710, Pakistan*
5. *Department of Applied Mathematics, Faculty of Mathematical Sciences, Shahrood University of Technology, Semnan, Iran*

***Corresponding Author:** Email: rmadah733@gmail.com

(Received 05 May 2023; accepted 15 Aug 2023)

Abstract

Background: Liver hepatocellular carcinoma (LIHC) is a common cancer with a poor prognosis and high recurrence rate. We aimed to identify potential biomarkers for LIHC by investigating the involvement of hub genes, microRNAs (miRNAs), transcription factors (TFs), and protein kinases (PKs) in its occurrence.

Methods: we conducted a bioinformatics analysis using microarray datasets, the TCGA-LIHC dataset, and text mining to identify differentially expressed genes (DEGs) associated with LIHC. They then performed functional enrichment analysis and gene-disease association analysis. The protein-protein interaction network of the genes was established, and hub genes were identified. The expression levels and survival analysis of these hub genes were evaluated, and their association with miRNAs, TFs, and PKs was assessed.

Results: The analysis identified 122 common genes involved in LIHC pathogenesis. Ten hub genes were filtered out, including *CDK1, CCNB1, CCNB2, CCNA2, ASPM, NCAPG, BIRC5, RRM2, KIF20A,* and *CENPF*. The expression level of all hub genes was confirmed, and high expression levels of all hub genes were correlated with poor overall survival of LIHC patients.

Conclusion: Identifying potential biomarkers for LIHC can aid in the design of targeted treatments and improve the survival of LIHC patients. The findings of this study provide a basis for further research in the field of LIHC and contribute to the understanding of its molecular pathogenesis.

Keywords: Hepatocellular carcinoma; Genes; Molecular pathway; Systems biology

Introduction

Hepatocellular Carcinoma (HCC) is one of the malignant diseases whose prevalence is increasing day by day. So far, in most cases, the etiology of the patient is related to some viruses, including HBV and HCV(1) . However, in addition to viruses, some other factors also play a role in the

pathogenesis of the disease(2). Disruption of gene expression as well as molecular pathways can also play a role in the pathogenesis of HCC (2).

Nowadays, various bioinformatics methods have been used to uncover novel biomarkers in different cancers to diagnose and treat (3, 4). The use of integrated bioinformatics analysis in cancer research has yielded considerable development.

The prognostic genes were investigated in HCC using microarray datasets and a bioinformatics approach. They found ten hub genes containing *CDC20, AURKA, FTCD, CKS1B, UBE2C, TOP2A, KIF20A, CCNB2, PTTG1,* and *CDKN3* correlate with the survival of patients with HCC (5).

The comprehensive identification of the main gene accountable for HCC's pathogenesis has not been accomplished. Although some studies have investigated genes and molecular pathways involved in HCC pathogenesis, most studies have focused on the diagnostic and prognostic value of genes (5).

In this study, we explored potential biomarkers associated with the occurrence and progression of HCC using a bioinformatics approach, employing three methods: TCGA, Meta-analysis, and text mining. The investigation comprehensively assessed primary hub genes, pathways, microRNAs (miRNAs), transcription factors (TFs), and protein kinases (PKs) involved in HCC. Furthermore, we evaluated and validated the relationship between these hub genes and the survival rate of patients.

Materials and Methods

TCGA Dataset and Screening DEGs

The raw RNA-Seq data were downloaded from the Cancer Genome Atlas (TCGA) database via TCGAbiolinks package of R using the following criteria: 422 cases (371 cancerous and 51 normal samples), project (TCGA-LIHC), data category (Transcriptome Profiling), data type (Gene Expression Quantification) and workflow type (STAR - Counts). The edgeR and limma packages of R software were used to normalize and screen differentially expressed genes (DEGs) between cancerous and normal samples. DEGs were screened |log2fold change| greater than 1, and adjusted *P*-value lower than 0.05.

Microarray Datasets and Screening DEGs

The datasets for gene expression analysis, namely GSE101685, GSE62232, and GSE45267, were obtained from the Gene Expression Omnibus (GEO) Database and their complete details are outlined in Table 1. All datasets were built on the GPL570 platform. To eliminate any batch effects in the three datasets, the SVA package of R software was employed for merging and removal purposes. The limma package was employed to detect the genes that showed differential expression between cancerous and normal samples. The DEGs were subsequently filtered based on a |log2fold change| that was greater than 1 and an adjusted *P*-value that was lower than 0.05.

Table 1: The full information of microarray datasets

Datasets	*Platform*	*Disease*	*Affected*	*Control*	*Sample*
GSE101685	GLP570	Hepatocellular carcinoma	24	8	Liver tissue
GSE62232	GLP570	Hepatocellular carcinoma	81	10	Liver tissue
GSE45267	GLP570	Hepatocellular carcinoma	48	39	Liver tissue

Identifying LIHC-related genes by using text mining

Text mining was done using the online tool GenCLip3 (6). GenCLip3 extracts the gene names associated with keywords from the PubMed literature. The following criteria were used: search term= Hepatocellular carcinoma and search years range= 1980-2020. The gene names that were in at least three publications were filtered.

Intersection of genes

For identifying overlapped genes between TCGA, microarray, and text mining analyses, Venn diagrams were applied. The shared genes between all analyses were used for upstream analysis.

Functional enrichment analysis of common genes

Gene ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis related to the overlapping genes were executed by using the GeneCodis software. Significant GO and KEGG terms were screened with a threshold *P*-value of 0.05.

Gene-disease association analysis

To evaluate the association between genes and human disease, the DisGeNET database was accessed via GeneCodis software, which is one of the most comprehensive databases of genes and variations linked to human disorders. A statistical value cutoff was set at an adjusted *P*-value lower than 0.05.

Protein-Protein interaction network establishment and hub genes identification

The STRING web tool, one of the largest online repositories of available protein-protein interactions (PPI), was used to predict interactions between proteins (7). A confidence scores greater than 0.7 was used to set the interaction parameter. PPI network was depicted via Cytoscape (8). In order to detect the key model of PPI, the Molecular complex detection (MCODE) plugin of Cytoscape was used through default parameters. Modules with MCODE scores > 5 and the number of nodes > 5 were regarded as significant modules (9). The cytoHubba plug-in of Cytoscape was used to determine hub genes through the degree method (10).

Validation of hub genes and survival analysis

In order to confirm the expression level of hub genes, Gene Expression Profiling Interactive Analysis 2 (GEPIA2) was used (11), which consists of RNA sequencing expression data from TCGA and the GTEx projects. The One-way ANOVA method and q-value lower than 0.05 were considered for differential expression analysis. Besides, the effect of hub genes on survival rate in HCC was evaluated through the Kaplan-Meier plotter (12). This web tool contains information about the impact of 54,000 genes on survival in different cancers. The hazard ratio (HR) and log-rank *P*-value were computed for this analysis.

Evaluation of the association between hub genes with miRNAs

The target miRNAs of hub genes were predicted by using the miRTarBase linked to Enrichr software (13). Statistical significance was regarded as an adjusted *P*-value lower than 0.05.

Evaluation of the association between hub genes with TFs and PKs

TFs and protein kinases PKs that potentially regulate the expression of hub genes were found using the X2K web tool. A hypergeometric *P*-value lower than 0.05 was considered statistically significant.

Results

TCGA, microarray, and text-mining analyses

Genes involved in HCC pathogenesis were investigated based on three methods: TCGA, Meta-analysis, and text mining. First, DEGs were acquired from the TCGA-LIHC dataset. Altogether 3640 DEGs, containing 1346 up- and 2294 down-regulated, were found. Then, a meta-analysis of the microarray datasets (GSE101685,

GSE62232, and GSE45267) was conducted, and 295 DEGs, including 94 up and 201 down-regulated genes, were found. Moreover. 2345 gene names were mined from literature using a text-miming approach. Altogether, 122 common genes involved in HCC pathogenesis were determined between all methods (Fig. 1).

Fig. 1: Common genes involved in HCC pathogenesis identified by TCGA analysis, meta-analysis of microarray datasets, and text mining; 122 common genes were found between all three methods, with 94 up-regulated and 28 down-regulated genes (Original)

Functional enrichment analysis of common genes

The GeneCodis software was employed to carry out a functional enrichment analysis on the genes. GO analysis displayed that in BP, the genes were mostly enriched for the GO terms such as "cell division", "cell cycle", and "epoxygenase P450 pathway". CC analysis showed that the genes were significantly enriched for the terms such as "spindle", "cytoplasm", and "midbody". For MF, the genes were enriched for the GO terms such as "oxidoreductase activity, acting on paired donors, with incorporation or reduction of molecular oxygen, reduced flavin or flavoprotein as one donor, and incorporation of one atom of oxygen", "aromatase activity" and "arachidonic acid epoxygenase activity". In addition, KEGG pathway analysis revealed that genes mainly enriched for the pathway terms such as "cell cycle", "caffeine metabolism" and "chemical carcinogenesis-DNA adducts" (Fig. 2).

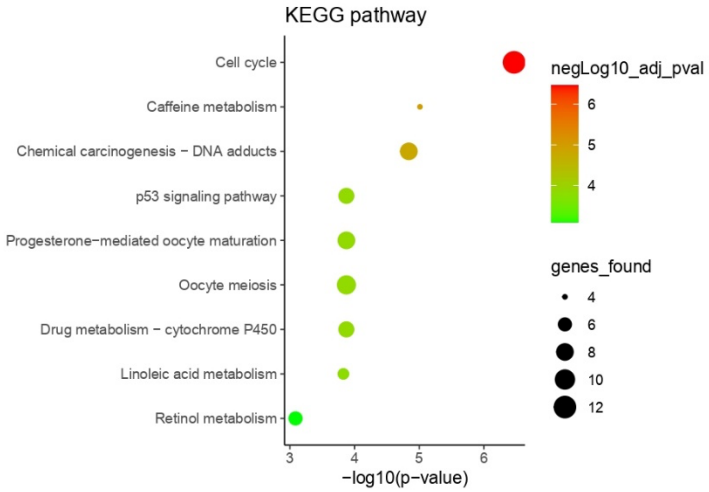

Fig. 2: Significant pathways associated with HCC pathogenesis identified by KEGG pathway analysis (Original)

Gene-disease association analysis

The correlation of genes with the human disease were assessed by the DisGeNET database linked to GeneCodis software. According to the results, genes mostly correlated with "Liver carcinoma", "Polycystic Ovary Syndrome" and "Sclerocystic Ovaries" (Fig. 3).

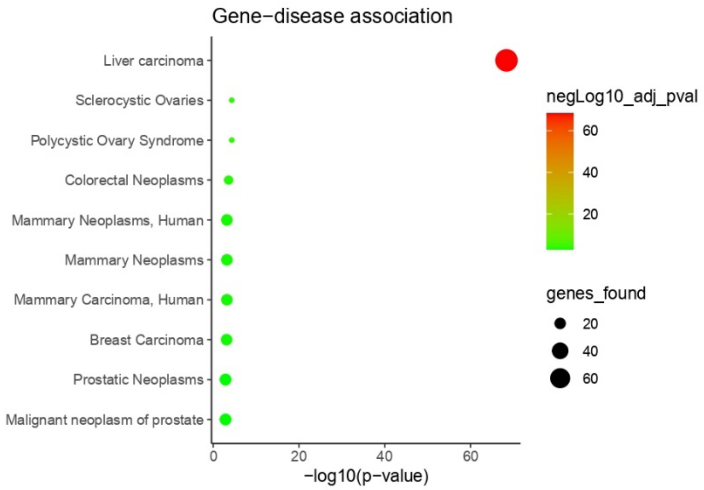

Fig. 3: Correlation of genes with various human diseases using the DisGeNET database (Original)

PPI construction analysis and selection of hub genes

The STRING database was used to construct the interactions network between proteins, and Cytoscape was employed to visualize interactions. The results revealed a comprehensive PPI network comprising 109 nodes and 999 edges (Fig. 4). Upon conducting the MCODE analysis, the investigation identified two highly significant modules, each constituting a key module within the PPI network (Fig. 5). The identification of key modules through MCODE analysis indicates specific clusters of proteins that might play crucial roles in regulating critical pathways or functional relationships within the network. The top 10 hub genes were chosen through the cytoHubba plugin. Based on the degree method, hub genes included *CDK1, CCNB1, CCNB2, CCNA2, ASPM, NCAPG, BIRC5, RRM2, KIF20A,* and *CENPF*. The topological information of hub genes is shown in Table 2.

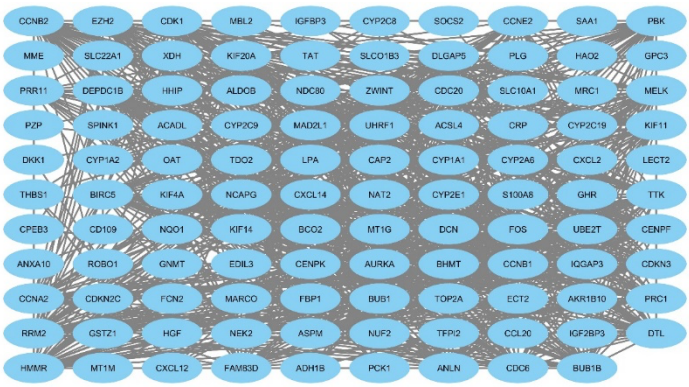

Fig. 4: PPI network of genes, where the blue ellipses represent genes and the gray lines indicate interactions between them. The network comprises 109 nodes and 999 edges. (original)

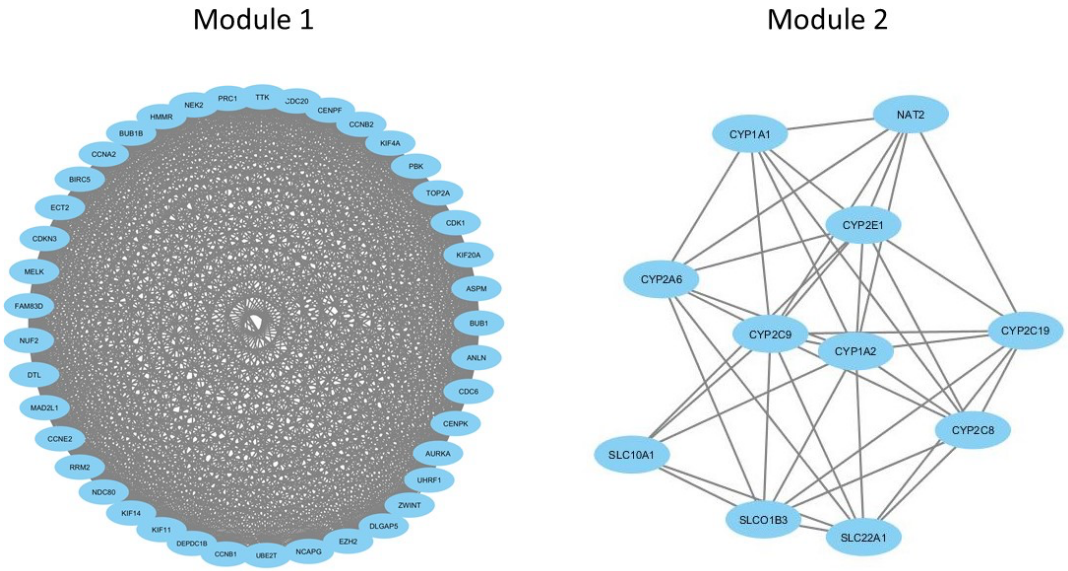

Fig. 5: Significant modules detected in the PPI network using the MCODE Plugin (module 1 with a score of 39.55 and module 2 with a score of 8). (original)

Table 2: The topological information of hub genes.

Name	*Degree*	*Betweenness*	*Closeness*	*Radiality*	*Stress*
CDK1	45	0.035769	0.418605	0.845679	3728
CCNB1	44	0.026914	0.416988	0.84465	3260
CCNB2	44	0.007227	0.388489	0.825103	880
ASPM	44	0.061777	0.407547	0.838477	2952
NCAPG	43	0.019325	0.387097	0.824074	1022
BIRC5	43	0.024353	0.415385	0.843621	2856
RRM2	43	0.078821	0.404494	0.83642	4162
KIF20A	43	0.012683	0.394161	0.829218	1526
CENPF	43	0.068916	0.415385	0.843621	4324

Validation of hub genes and survival analysis

To verify the expression level for hub genes, GEPIA2 was used. Results indicated that the expression level of all hub genes significantly up-regulated and validated (Fig. 6). Moreover, to identify the prognostic value of hub genes, the overall survival analysis of the hub genes was plotted using the Kaplan-Meier plotter. The curves showed that overexpression of all hub genes is significantly associated with decreased overall survival times of patients with HCC (Fig. 7).

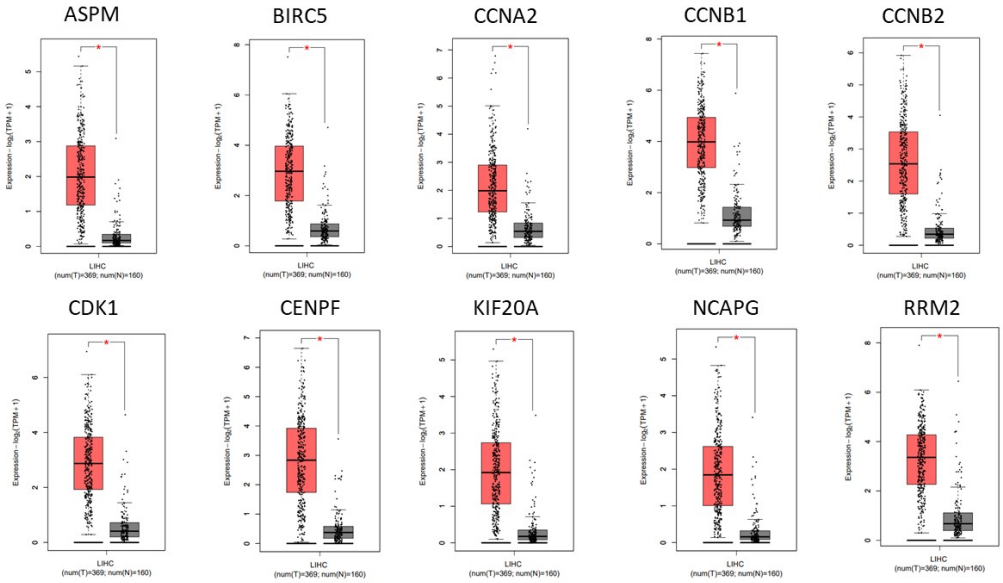

Fig. 6: The results of the validation analysis of 10 hub genes using GEPIA2 web tool. This figure shows a comparison of mRNA expression levels between HCC tissues (red boxes) and control tissues (gray boxes) for all hub genes (Original)

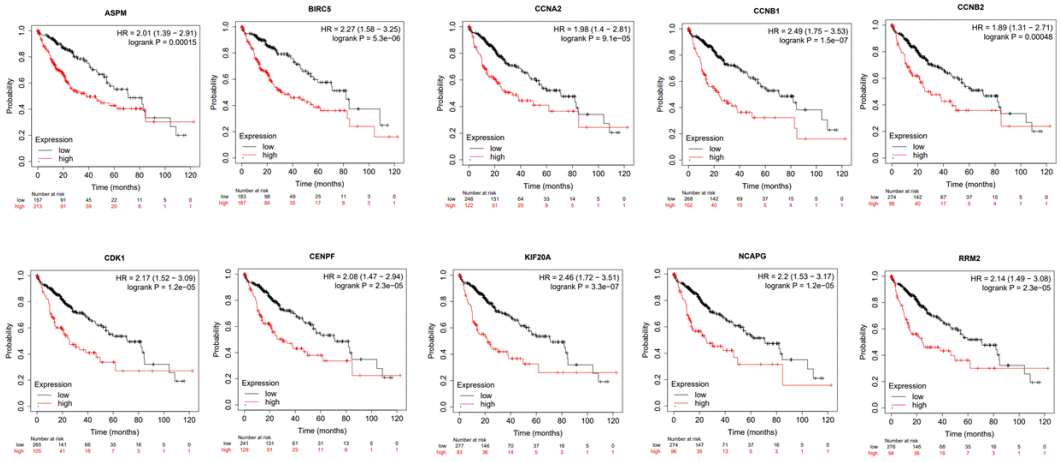

Fig. 7: Kaplan-Meier survival plots of hub genes using Kaplan-Meier plotter software, where the black and red lines indicate low- and high-risk patient groups. The curves demonstrate a significant association between overexpression of all hub genes and decreased overall survival times in patients with HCC (Original)

Interaction of hub genes with miRNAs

In order to evaluate the interaction of hub genes with miRNAs, the miRTarBase linked to Enrichr software was used. Hub genes interacted with some miRNAs, which played a role in HCC pathogenesis. These miRNAs, including has-miR-193b-3p, has-miR-205-5p, and has-let-7b-5p, were considered potential critical miRNAs.

Interaction of hub genes with TFs and PKs

In order to evaluate the interaction of hub genes with TFs and PKs, the X2K online tool was used. The most important TFs included FOXM1 and SIN3A, which interact with most of the hub genes. In addition, the most important PKs in-

cluded MAPK14 and CSNK2A1, which interact with most of the hub genes.

Discussion

In the bioinformatics analysis, the expression of some genes, such as *CCNB1, CDK1,* and *CCNB2*, increased in hepatocellular carcinoma (14). Reduced expression of *CCNB2, CCNB1,* and *CDK1* stop cells in the G0/G1 of the cell cycle and cause the induction of apoptosis and inhibition of cell invasion by blocking the p53 signaling pathway. Also, the reduction of CDK1/CCNB1 expression can lead to the increase of p53 and p21Cip1 protein expression as tumor suppressors in cancer cells (14). Decreased expression of miR-199a-3p is significantly observed in HCC cells (14). miR-199a-3p impedes the expression of CCNB1 and CDK1 through activating the p53 signaling pathway, thereby leading to apoptosis and inhibition of cell invasion (14). The mentioned genes are prognostic biomarkers and are related to inflammatory cell infiltration and response to immunotherapy in HCC (15). CNA (copy number alteration) of CDK1 and CCNB1 genes have a significant relationship with the infiltration of immune cells in the liver tissue of HCC patients (15).

Increased expression of the *CCNA2* gene leads to the induction of cell proliferation and tumorigenesis through the activation of the FXR-miR-22-CCNA2 signaling pathway. These results may be helpful in identifying diagnostic and prognostic markers of HCC patients (16). Overexpression of ASPM is seen in malignant liver cells and correlates with poor prognosis. ASPM interacts with disheveled-2 (Dvl2) and inhibits Dvl2 degradation, and leads to the increase of Dvl2 protein by activation of Wnt/β-catenin signaling pathway in HCC, so that disarrangement of the Wnt/ASPM/Dvl2-β-catenin signaling pathway may have clinical worth in management of patients with HCC (17). Elevated expression of NCAPG (non-SMC condensing I complex subunit G) correlates with poor prognosis in HCC patients and is found as a prognostic factor, leading to cell proliferation through activation of PI3K/AKT/FOXO4 and p53 signaling pathways. In addition to NCAPG expression status, disease stage (T AJCC) is also an independent predictor of overall survival (OS) (18). NCAPG expression level correlates with hypomethylation status and has a positive correlation with tumor immune cell infiltration (18). The PI3K/AKT signaling pathway is one of the most critical pathways for the development of HCC. Disturbance in this pathway leads to decreased cell growth and increased apoptosis (19). The level of expression of miR-181c in HCC cells is notably lower than that in normal hepatocytes. There is a reverse correlation between the expression of miR-181c and NCAPG in HCC (20). Its lower expression stops cell proliferation by inhibiting the Notch 2 signaling pathway (20). Overexpression of BIRC5 (survivin) has been observed in HCC and is correlated with poor outcomes (21). Survivin contributes to carcinogenesis through different mechanisms, including interactivity with caspase 3 and 7, Bax and Fas induced suppression of apoptosis, cytokines and cell cycle progression control, and contribution in a diversity of signaling pathways such as p53, Wnt, hypoxia, transforming growth factor (TGF) and the Notch signaling pathway (22). Casein kinase 2 alpha protein 1 (CSNK2A1) leads to the proliferation of malignant cells, invasion, distal metastasis, and resistance to chemotherapeutic agents through activating the Wnt signaling pathway (23). Ribonucleotide reductase subunit M2 (RRM2) leads to the deoxynucleotide triphosphates production, which is crucial for DNA biosynthesis (24). RRM2 plays a vital role in tumor proliferation, and over-expression of RRM2 correlates with poor disease outcomes. A higher level of RRM2 protein expression can be a helpful biomarker for anticipating early recurrence after therapeutic surgery (24). Inhibition of RRM2 notably decreases the expression of the antiapoptotic protein Bcl-2 signaling pathway (24). The genes, including CDK1, RRM2, ASPM, and PBK may be essential for the evolution of HCC from cirrhosis through activating the p53 signaling pathway (25). Therefore, the p53 signaling pathway is cru-

cial in the progression of HCC (25). Gli2 (Glioma-associated oncogene 2), an early transcriptional regulator of the Hedgehog (Hh) signaling pathway, is important for HCC (26). Lower expression of Gli2 leads to the suppression of malignant cell proliferation, and its overexpression leads to the increase in the expression of KIF20A through activating Forkhead Box M1 (FoxM1)-MMB complex-mediated transcription (26). KIF20A is a new downstream target of the Hh signaling pathway, which is essential for tumor proliferation (26, 27). CENPF (Centromere Protein F) may be used as a potential prognostic biomarker and novel therapeutic target for HCC (28). Overexpression of CENPF through the MAPK signaling pathway results in the progression of HCC and abnormal cell division within the cell cycle through the CENPF-E2F1/CDK1 signaling pathway (28). Thus, genes interacting with CENPF and their associated signaling pathways may give indications for the development of targeted CENPF-mediated therapy (28).

Reduced expression of miR-205-5p is found in HCC cell lines, resulting in resistance to chemotherapeutic agents through the PTEN/JNK/ANXA3 pathway (29). Treatment with 5-fluorouracil (5-Fu) enhances the expression of miR-205-5p in HCC malignant cells. The PTEN/JNK/ANXA3 signaling pathway alterations are linked to these consequences (29). Overexpression of small nucleolar RNA host gene 16 (SNHG16) occurs in HCC and indicates poor outcomes in HCC patients. The NHG16 facilitates the transition of the G2/M cell cycle by directly affecting the let-7b-5p/CDC25B/CDK1 signaling path, and it leads to distant metastasis and (Epithelial-mesenchymal transition) through the let-7b-5p/HMGA2 axis. These findings may provide a novel approach for treating HCC and its molecular pathogenesis. Transcription factor HMGA2 is involved in tumor proliferation and EMT progression by activating numerous signaling pathways, including TGF β1/Smad3, MAPK, and Wnt/β-catenin (30). FOXM1 is a transcription factor that plays a significant role in cancer development and progression, including HCC (31). In HCC patients, the increased expression of FOXM1 and CCNB1 is strongly linked to poor prognosis, and they are expressed simultaneously in HCC. In HCC cell lines, the overturning of FOXM1 significantly hampered the level of expression of CCNB1 at mRNA and protein levels (31). It actually binds straight to the CCNB1 promoter region and regulates CCNB1 gene expression levels at the transcriptional level (31). CCNB1 is critical to FOXM1-induced growth in HCC cells. FOXM1 is engaged in numerous oncogenic pathways in HCC, like YAP1 and Ras-ERK1/2 (31). FOXM1, TCF7L1, E2F4, and SIN3A are the main TFs for HCC (32). overexpression of miR-210 causes suppression of proliferation and induction of apoptosis through targeting SIN3A, a transcriptional regulator. Changes in signaling pathways like Ras/MAPK, Wnt/β-catenin, and retinoblastoma pathways occur in HCC cells (32). E2F4 is one of the crucial members of TFs that participate in innate immune responses like the expression of CD14 and toll-like receptor 8 (TLR8) and downstream signal transducer and activator of transcription (STAT1). Overexpression of E2F4 in HCC tumor tissue versus adjacent normal tissue. Its increased expression is significantly associated with poor prognosis in HCC patients (33).

This study provides a comprehensive analysis of potential biomarkers for HCC, employing a combination of bioinformatics approaches. However, the study relies on data obtained from public databases, and the identified biomarkers necessitate further validation through experimental studies. Notably, the study did not account for other factors that could affect the prognosis of HCC patients, such as tumor size, grade, and stage. Moreover, potential confounding factors like comorbidities and environmental exposures were not considered, which might influence the expression of the identified biomarkers. Additional validation studies are essential to establish the clinical utility of the identified biomarkers and their potential application in personalized treatment for HCC patients.

Conclusion

Finding out the signaling pathways, critical genes, miRNAs, TFs, and PKs involved in the growth of hepatocellular carcinoma can be used in designing targeted therapy and increasing the overall survival of HCCs. Besides, investigating the increase or decrease in the expression of genes and miRNAs and signaling pathways in HCC is very useful because it assists in identifying diagnostic and prognostic biomarkers. Although, additional examinations are required to evaluate these biomarkers.

Journalism Ethics considerations

Ethical issues (Including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, redundancy, etc.) have been completely observed by the authors.

Availability of data and materials

The datasets used in this research are available in GEO and TCGA databases.

Funding

No funding was received.

Conflict of interests

All authors declare that they have no competing interests.

References

1. Dimri M, Satyanarayana A (2020). Molecular signaling pathways and therapeutic targets in hepatocellular carcinoma. *Cancers (Basel)*, 12 (2): 491.
2. Khalaf AM, Fuentes D, Morshid AI, et al (2018). Role of Wnt/β-catenin signaling in hepatocellular carcinoma, pathogenesis, and clinical significance. *J Hepatocell Carcinoma*, 5:61-73.
3. Hanauer DA, Rhodes DR, Sinha-Kumar C, Chinnaiyan AM (2007). Bioinformatics approaches in the study of cancer. *Curr Mol Med*, 7 (1): 133-41.
4. Banwait JK, Bastola DR (2015). Contribution of bioinformatics prediction in microRNA-based cancer therapeutics. *Adv Drug Deliv Rev*, 81: 94-103.
5. Xu L, Tong T, Wang Z, et al (2020). Identification of hub genes and analysis of prognostic values in hepatocellular carcinoma by bioinformatics analysis. *Am J Med Sci*, 359 (4): 226-34.
6. Wang JH, Zhao LF, Wang HF, et al (2019). GenCLiP 3: mining human genes' functions and regulatory networks from PubMed based on co-occurrences and natural language processing. *Bioinformatics*, btz807.
7. Szklarczyk D, Gable AL, Nastou KC, et al (2021). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucleic Acids Res*, 49 (D1): D605-D12.
8. Shannon P, Markiel A, Ozier O, et al (2003). Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res*, 13 (11): 2498-504.
9. Bader GD, Hogue CW (2003). An automated method for finding molecular complexes in large protein interaction networks. *BMC Bioinformatics*, 4:2.
10. Chin C-H, Chen S-H, Wu H-H, et al (2014). cytoHubba: identifying hub objects and sub-networks from complex interactome. *BMC Syst Biol*, 8 Suppl 4(Suppl 4):S11.
11. Tang Z, Kang B, Li C, et al (2019). GEPIA2: an enhanced web server for large-scale expression profiling and interactive analysis. *Nucleic Acids Res*, 47 (W1): W556-W560.
12. Lánczky A, Győrffy B (2021). Web-based survival analysis tool tailored for medical research (KMplot): development and implementation. *J Med Internet Res*, 23 (7): e27633.
13. Kuleshov MV, Jones MR, Rouillard AD, et al (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res*, 44 (W1): W90-W7.

14. Jin J, Xu H, Li W, et al (2020). LINC00346 acts as a competing endogenous RNA regulating development of hepatocellular carcinoma via modulating CDK1/CCNB1 axis. *Front Bioeng Biotechnol*, 8: 54.
15. Zou Y, Ruan S, Jin L, et al (2020). CDK1, CCNB1, and CCNB2 are prognostic biomarkers and correlated with immune infiltration in hepatocellular carcinoma. *Med Sci Monit*, 26: e925289.
16. Zhang H, Liu R, Sun L, et al (2021). Comprehensive analysis of gene expression changes and validation in hepatocellular carcinoma. *Onco Targets Ther*, 14: 1021-1031.
17. Zhang H, Yang X, Zhu L, et al (2021). ASPM promotes hepatocellular carcinoma progression by activating Wnt/β-catenin signaling through antagonizing autophagy-mediated Dvl2 degradation. *FEBS Open Bio*, 11 (10): 2784-99.
18. Guo Z, Zhu Z (2021). NCAPG is a prognostic biomarker associated with vascular invasion in hepatocellular carcinoma. *Eur Rev Med Pharmacol Sci*, 25(23):7238-51.
19. Gong C, Ai J, Fan Y, et al (2019). NCAPG promotes the proliferation of hepatocellular carcinoma through PI3K/AKT signaling. *Onco Targets Ther*, 12:8537-8552.
20. Ai J, Gong C, Wu J, et al (2019). MicroRNA-181c suppresses growth and metastasis of hepatocellular carcinoma by modulating NCAPG. *Cancer Manag Res*, 11: 3455-3467.
21. Xu R, Lin L, Zhang B, et al (2021). Identification of prognostic markers for hepatocellular carcinoma based on the epithelial-mesenchymal transition-related gene BIRC5. *BMC Cancer*, 21 (1): 687.
22. Chen X, Duan N, Zhang C, et al (2016). Survivin and tumorigenesis: molecular mechanisms and therapeutic strategies. *J Cancer*, 7 (3): 314-23.
23. Wu R, Tang W, Qiu K, et al (2021). An Integrative Pan-Cancer Analysis of the Prognostic and Immunological Role of Casein Kinase 2 Alpha Protein 1 (CSNK2A1) in Human Cancers: A Study Based on Bioinformatics and Immunohistochemical Analysis. *Int J Gen Med*, 14: 6215-6232.
24. Lee B, Ha SY, Song DH, et al (2014). High expression of ribonucleotide reductase subunit M2 correlates with poor prognosis of hepatocellular carcinoma. *Gut Liver*, 8 (6): 662-8.
25. He B, Yin J, Gong S, et al (2017). Bioinformatics analysis of key genes and pathways for hepatocellular carcinoma transformed from cirrhosis. *Medicine (Baltimore)*, 96 (25): e6938.
26. Xiong M, Zhuang K, Luo Y, et al (2019). KIF20A promotes cellular malignant behavior and enhances resistance to chemotherapy in colorectal cancer through regulation of the JAK/STAT3 signaling pathway. *Aging (Albany NY)*, 11 (24): 11905-11921.
27. Lu M, Huang X, Chen Y, et al (2018). Aberrant KIF20A expression might independently predict poor overall survival and recurrence-free survival of hepatocellular carcinoma. *IUBMB Life*, 70 (4): 328-35.
28. Huang Y, Chen X, Wang L, et al (2021). Centromere protein F (CENPF) serves as a potential prognostic biomarker and target for human hepatocellular carcinoma. *J Cancer*, 12 (10): 2933-2951.
29. Shao P, Qu W-K, Wang C-Y, et al (2017). MicroRNA-205-5p regulates the chemotherapeutic resistance of hepatocellular carcinoma cells by targeting PTEN/JNK/ANXA3 pathway. *Am J Transl Res*, 9 (9): 4300-4307.
30. Li S, Peng F, Ning Y, et al (2020). SNHG16 as the miRNA let-7b-5p sponge facilitates the G2/M and epithelial-mesenchymal transition by regulating CDC25B and HMGA2 expression in hepatocellular carcinoma. *J Cell Biochem*, 121 (3): 2543-58.
31. Chai N, Xie H-h, Yin JP, et al (2018). FOXM1 promotes proliferation in human hepatocellular carcinoma cells by transcriptional activation of CCNB1. *Biochem Biophys Res Commun*, 500 (4): 924-9.
32. Wang J, Tian Y, Chen H, et al (2018). Key signaling pathways, genes and transcription factors associated with hepatocellular carcinoma. *Mol Med Rep*, 17 (6): 8153-60.
33. Zheng Q, Fu Q, Xu J, et al (2021). Transcription factor E2F4 is an indicator of poor prognosis and is related to immune infiltration in hepatocellular carcinoma. *J Cancer*, 12 (6): 1792-1803.