Identification of Novel Biomarkers in Hepatocellular Carcinoma by Integrated Bioinformatical Analysis and Experimental Validation

Chen Liao
Yunnan University of Traditional Chinese Medicine
Lanlan Wang
Shaanxi University of Chinese Medicine
Xiaoqiang Li
Fourth Military Medical University Department of Social Sciences: Air Force Medical University
Jinyu Bai
Yunnan University of Traditional Chinese Medicine
Jieqiong Wu
Shaanxi University of Chinese Medicine
Wei Zhang
Shaanxi University of Chinese Medicine
Hailong Shi
Shaanxi University of Chinese Medicine
Xuesong Feng
Shaanxi University of Chinese Medicine
Xu Chao (chaoxu@sntcm.edu.cn)
Shaanxi University of Chinese Medicine https://orcid.org/0000-0001-5520-4834

Research

Keywords: Hepatocellular carcinoma, novel biomarkers, candidate small molecules, prognosis, bioinformatics analysis

Posted Date: June 16th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-533830/v1

Abstract

Background: Hepatocellular carcinoma (HCC) is one of the most common poorly prognosed virulent neoplasms of the digestive system. In this study, we identified novel biomarkers associated with the pathogenesis of HCC aiming to provide new diagnostic and therapeutic approaches for HCC.

Methods: Gene expression profiles of GSE62232, GSE84402,GSE121248 and GSE45267 were obtained in GEO database. Differential expressed genes (DEGs) between HCC and normal samples were identified using the GEO2R tool and Venn diagram software.Database for Annotation, Visualization and Integrated Discovery (DAVID) were used to carry out enrichment analysis on gene ontology (GO) and the Kyoto Encyclopaedia of Genes and Genomes pathway (KEGG). The protein-protein interaction (PPI) network of DEGs was constructed by the Search Tool for the Retrieval of Interacting Genes (STRING) and visualized by Cytoscape. Expressions and prognostic values of hub genes were validated through Kaplan-Meier plotter, Gene Expression Profiling Interactive Analysis (GEPIA), the Human Protein Atlas Database (HPA), western blot (WB) and quantitative real-time polymerase chain reaction (qRT-PCR). Additionally, potential small molecule drugs were screened by Connectivity Map (CMAP).

Results: A total of 100 overlapped DEGs were detected and results showed 23 of which were up-regulated with the rest being down-regulated. STRING screened the 70 edges and the 199 nodes in the PPI network. Survival analysis showed that aberrant mRNA expression of TOP2A, DTL, ANLN, CDKN3, BUB1B, CDK1, PBK, RRM2, RACGAP1, PRC1, NEK2, ECT2, CCNB1, HMMR, ASPM was significantly associated with a low survival rate. Results of WB and qRT-PCR showed that the expression levels of ANLN, CCNB1, DTL, RACGAP1, RRM2 and TOP2A were all increased in HCC tissues. Furthermore, CMAP predict suggest the 10 most vital small molecule drugs could reverse the progression of HCC.

Conclusions: Core DEGs (ANLN, CCNB1, DTL, RACGAP1, RRM2 and TOP2A) with poor prognosis and candidate drugs for HCC treatment were identified through integrated bioinformatic analysis.This study will contribute to providing prognostic biomarker and therapeutic strategies in HCC. Background : Hepatocellular carcinoma (HCC) is one of the most common poorly prognosed virulent neoplasms of the digestive system. In this study, we identified novel biomarkers associated with the pathogenesis of HCC aiming to provide new diagnostic and therapeutic approaches for HCC. Methods : Gene expression profiles of GSE62232, GSE84402,GSE121248 and GSE45267 were obtained in GEO database. Differential expressed genes (DEGs) between HCC and normal samples were identified using the GEO2R tool and Venn diagram software.Database for Annotation, Visualization and Integrated Discovery (DAVID) were used to carry out enrichment analysis on gene ontology (GO) and the Kyoto Encyclopaedia of Genes and Genomes pathway (KEGG). The protein-protein interaction (PPI) network of DEGs was constructed by the Search Tool for the Retrieval of Interacting Genes (STRING) and visualized by Cytoscape. Expressions and prognostic values of hub genes were validated through Kaplan-Meier plotter, Gene Expression Profiling Interactive Analysis (GEPIA), the Human Protein Atlas Database (HPA), western blot (WB) and quantitative real-time polymerase chain reaction (qRT-PCR). Additionally, potential small molecule drugs were screened by Connectivity Map (CMAP). Results : A total of 100 overlapped DEGs were detected and

results showed 23 of which were up-regulated with the rest being down-regulated. STRING screened the 70 edges and the 199 nodes in the PPI network. Survival analysis showed that aberrant mRNA expression of TOP2A, DTL, ANLN, CDKN3, BUB1B, CDK1, PBK, RRM2, RACGAP1, PRC1, NEK2, ECT2, CCNB1, HMMR, ASPM was significantly associated with a low survival rate. Results of WB and qRT-PCR showed that the expression levels of ANLN, CCNB1, DTL, RACGAP1, RRM2 and TOP2A were all increased in HCC tissues. Furthermore, CMAP predict suggest the 10 most vital small molecule drugs could reverse the progression of HCC. Conclusions : Core DEGs (ANLN, CCNB1, DTL, RACGAP1, RRM2 and TOP2A) with poor prognosis and candidate drugs for HCC treatment were identified through integrated bioinformatic analysis.This study will contribute to providing prognostic biomarker and therapeutic strategies in HCC.

Full Text

This preprint is available for download as a PDF.

Tables

Due to technical limitations, table 1 to 3 is only available as a download in the Supplemental Files section.

Figures

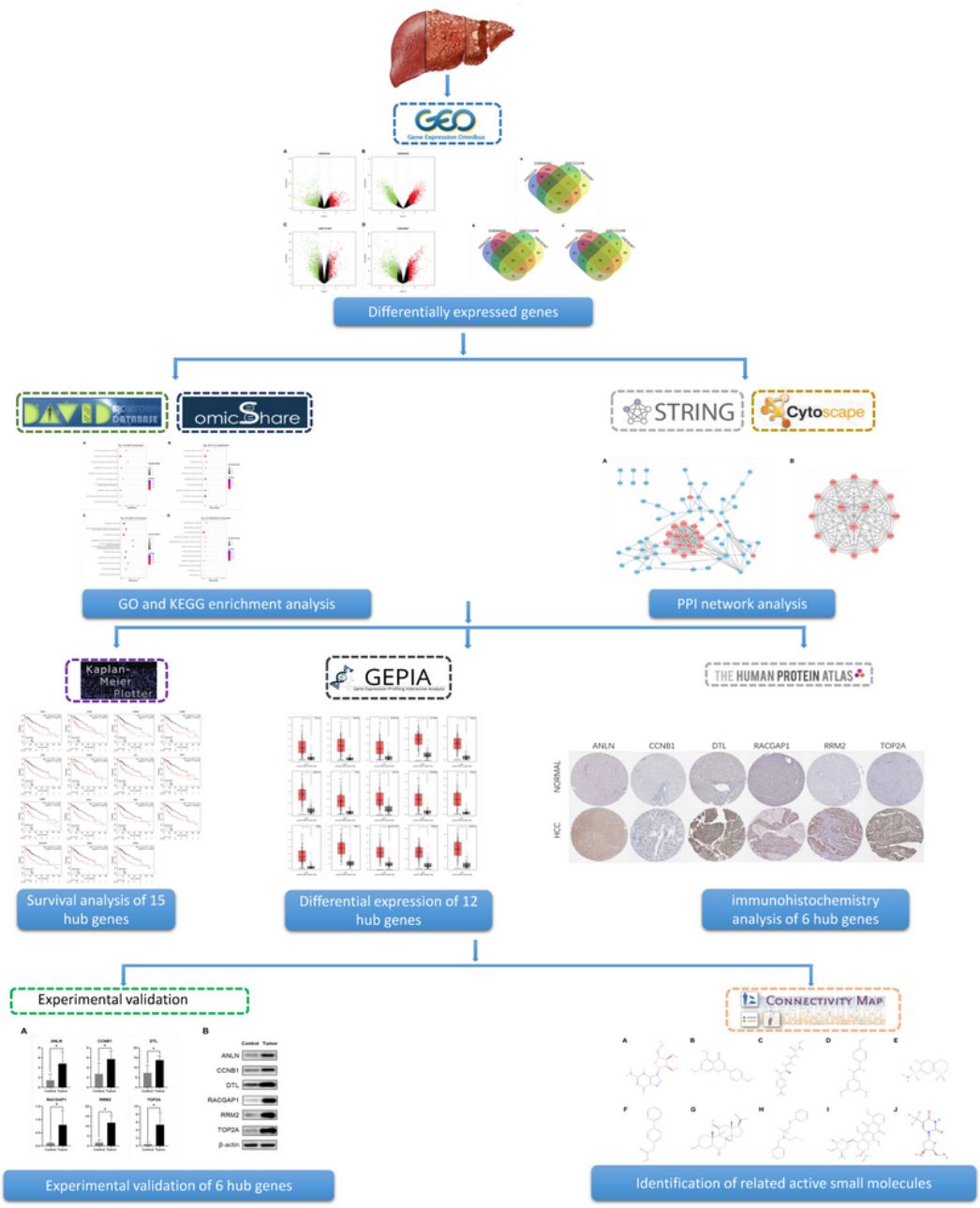

Figure 1

The workflow of this study

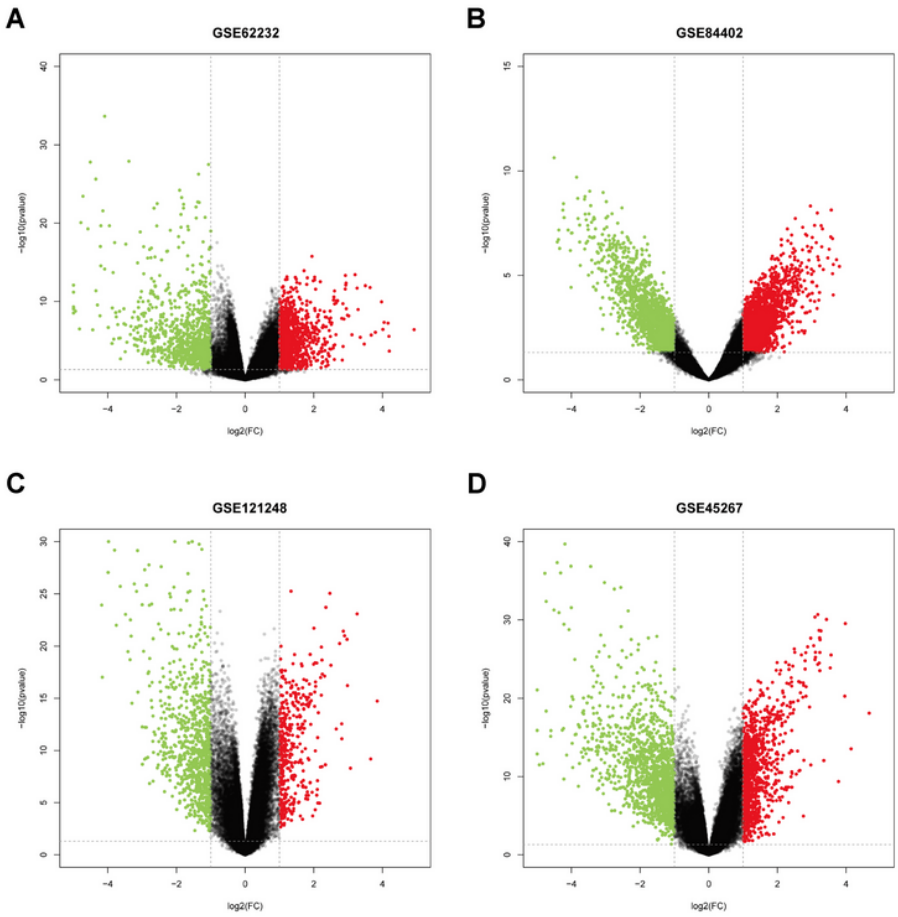

Figure 2

Volcano plot of gene expression profile data between HCC and normal tissues in each dataset. Red dots: significantly up-regulated genes in HCC; green dots: significantly down-regulated genes in HCC; black dots: non differentially expressed genes. P□0.05 and |log2FC|≥2 were considered as significant

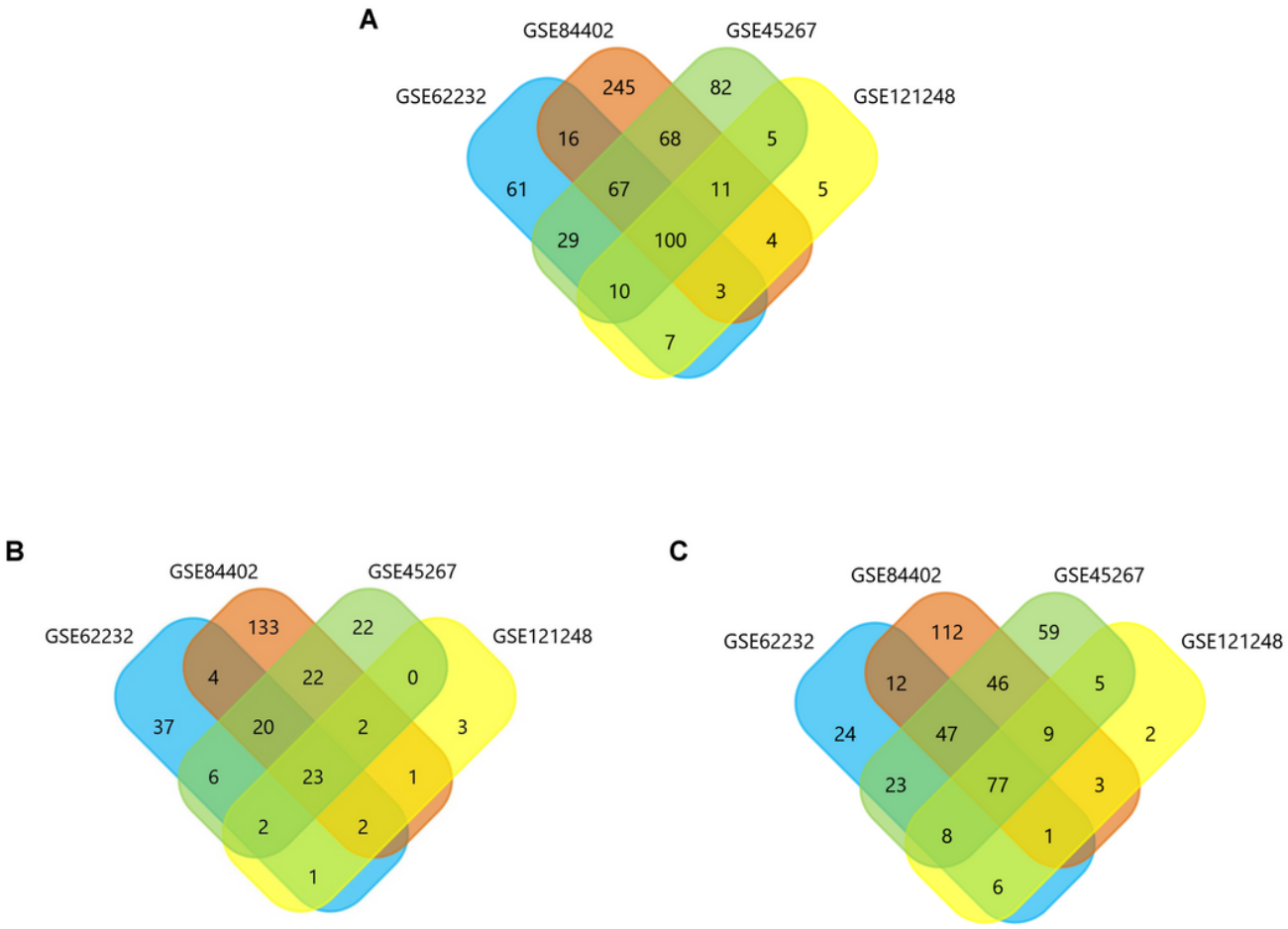

Figure 3

(A)Venn diagram of 100 overlapping DEGs from GSE62232, GSE84402, GSE121248 and GSE45267 datasets. (B) Up-regulated DEGs.　C) Down-regulated DEGs.

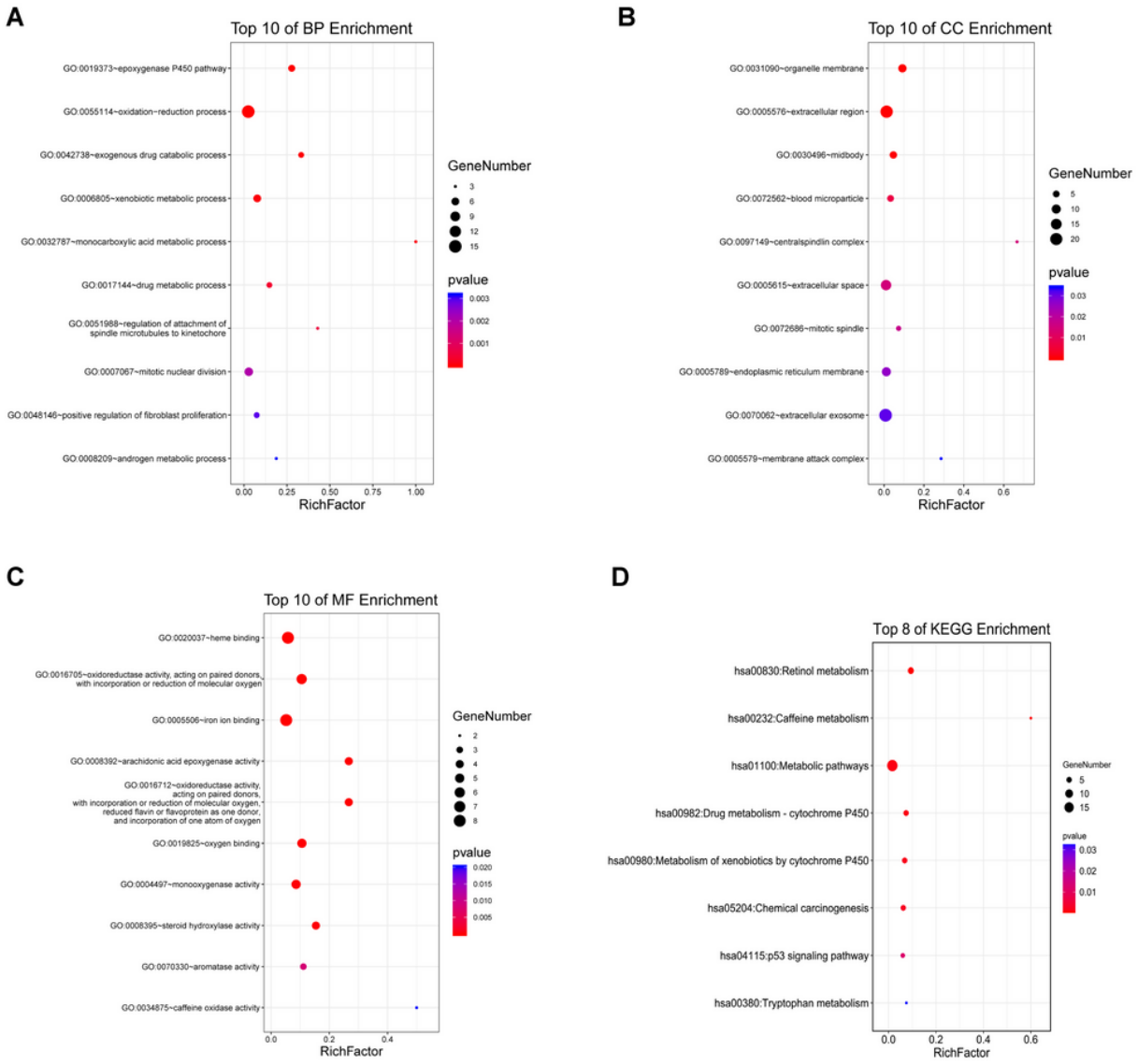

Figure 4

Functional and signaling pathway analysis of the overlapped DEGs in HCC. (A) Biological process. (B) Cellular components. (C) Molecular function and (D)KEGG pathway.

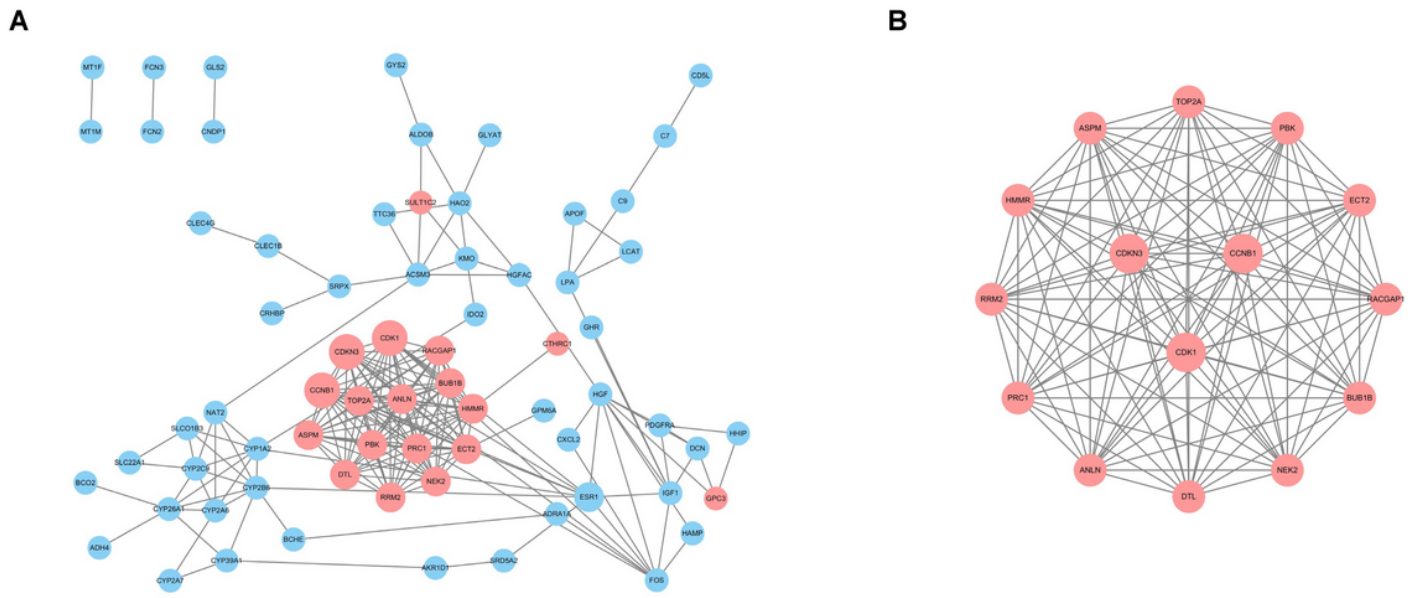

Figure 5

PPI network of the potential interactions between the overlapping DEGs.

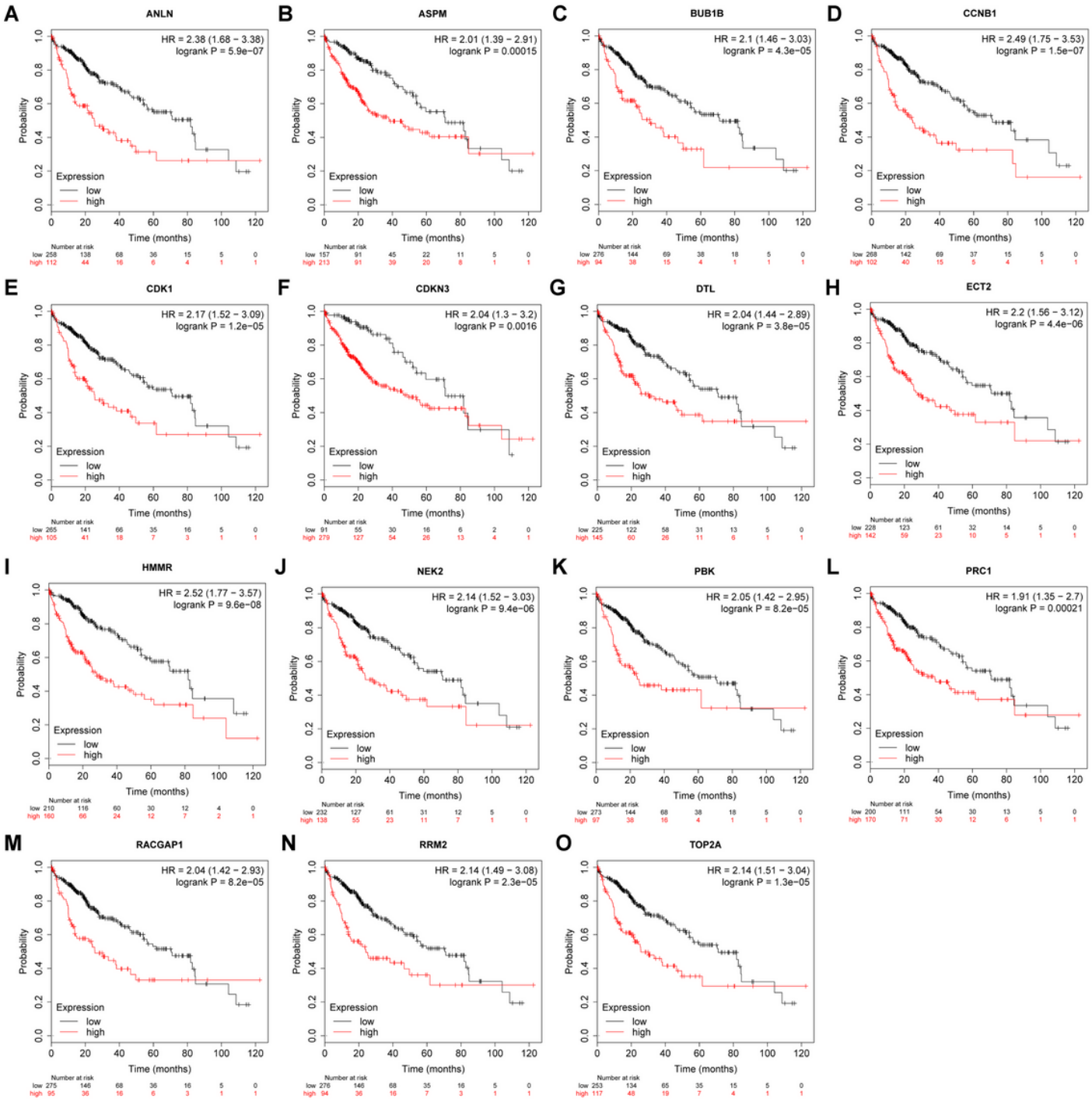

Figure 6

The prognostic information of the 15 core genes. Kaplan-Meier plotter online tools were used to identify the prognostic information of the 15 core genes and these genes had a significantly worse survival rate ($p < 0.05$).

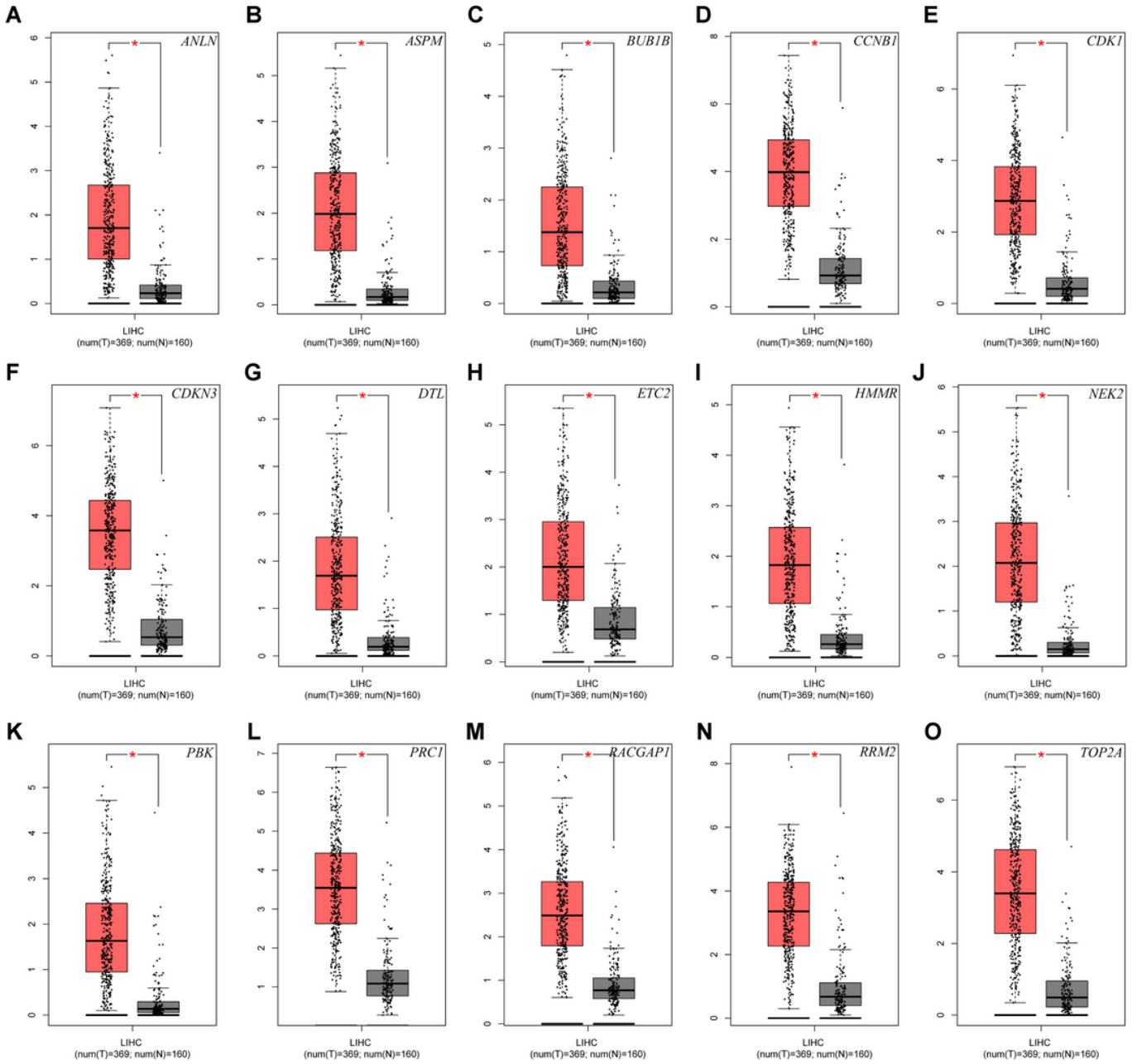

Figure 7

Significantly expressed 15 genes in HCC patients were analyzed by GEPIA website ($p<0.05$). Red color represents tumor samples and grey color represents normal samples.

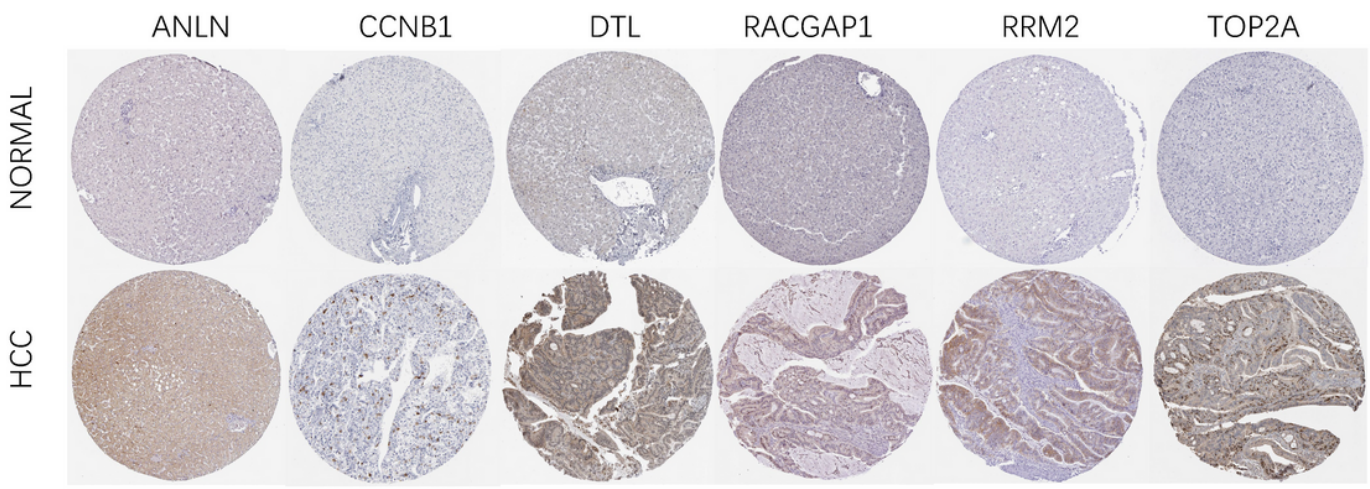

Figure 8

Representative immunohistochemistry staining results reveal the protein expression of hub genes in HCC and normal tissues.

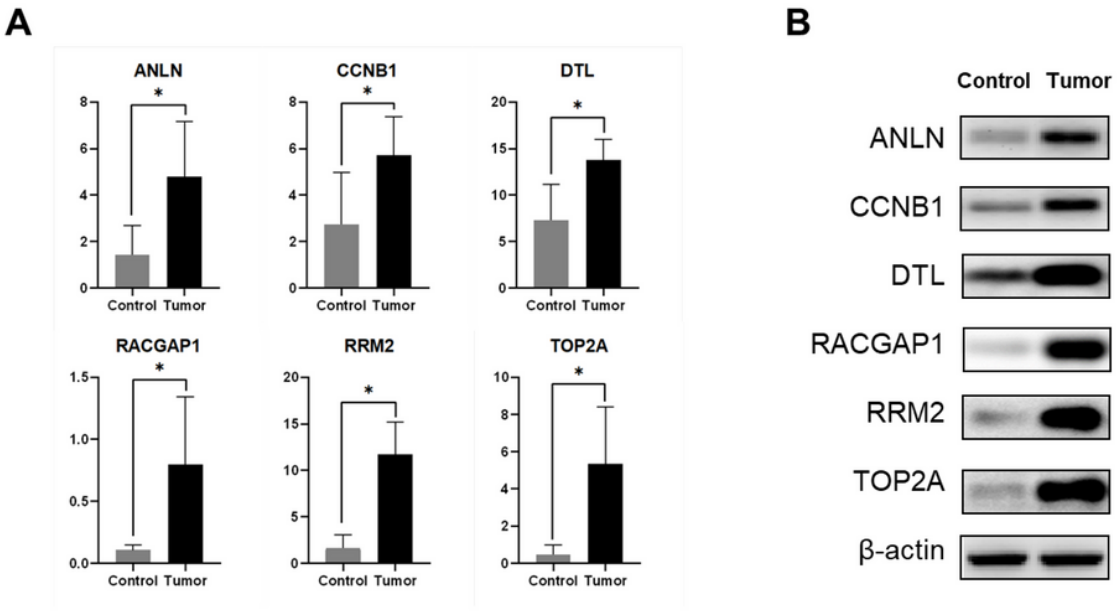

Figure 9

(A) qRT-PCR validation of these six hub genes in 3 paired HCC samples. (B) WB validation of these six hub genes in 3 paired HCC samples.

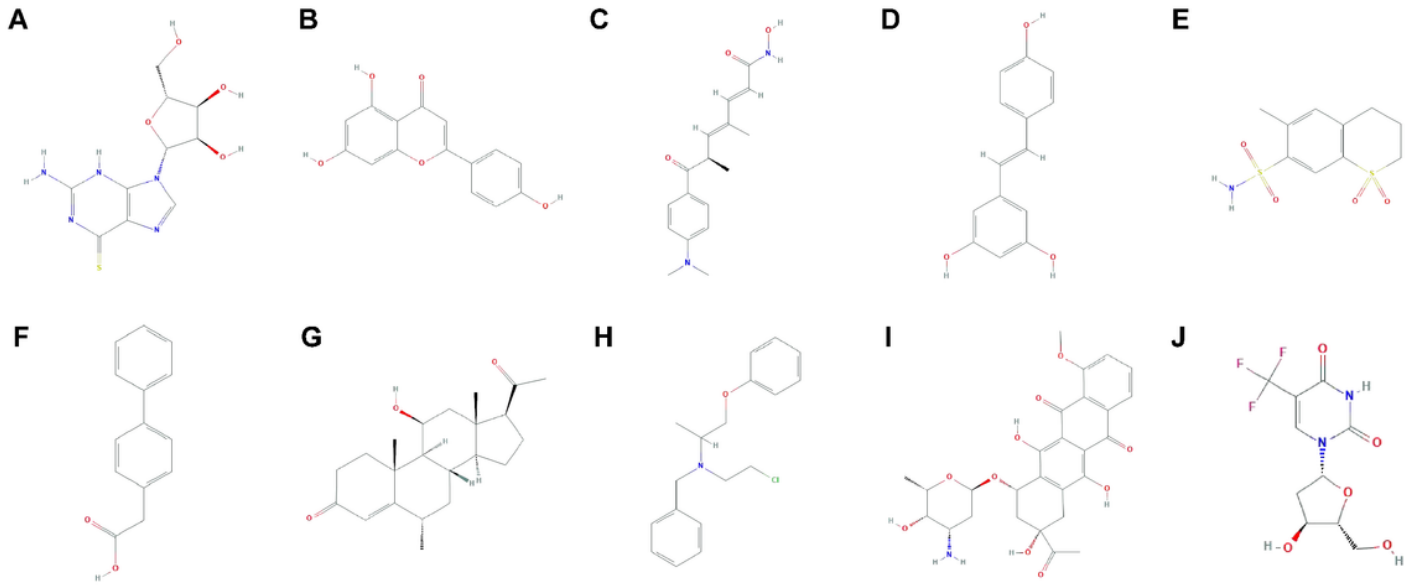

Figure 10

Chemical structures of the top 10 most significant molecules for HCC. (A) thioguanosine. (B apigenin. (C) trichostatin A. (D) resveratrol. (E) meticrane (F) felbinac. (G) medrysone. (H) phenoxybenzamine. (i) daunorubicin. (j) trifluridine.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Tables.pdf
- Supplementarymaterial.pdf