Original Article

Identification of genes predicting unfavorable prognosis in hepatitis B virus-associated hepatocellular carcinoma

Meng Sha[#], Jie Cao[#], Zhi-Peng Zong[#], Ning Xu, Jian-Jun Zhang, Ying Tong, Qiang Xia

Department of Liver Surgery, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai, China

Contributions: (I) Conception and design: M Sha, ZP Zong; (II) Administrative support: J Cao; (III) Provision of study materials or patients: N Xu; (IV) Collection and assembly of data: M Sha, J Cao; (V) Data analysis and interpretation: ZP Zong, JJ Zhang, Y Tong; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Ying Tong; Qiang Xia. Department of Liver Surgery, Renji Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200127, China. Email: tongyingsjtu@163.com; xiaqiang@shsmu.edu.cn.

Background: To identify potential key genes predicting unfavorable prognosis in hepatitis B virus (HBV)-associated hepatocellular carcinoma (HCC).

Methods: Gene expression profiles of GSE121248, GSE62232, and GSE55092 from the GEO database were obtained and analyzed. Differentially expressed genes (DEGs) between HBV-associated HCC tissues and adjacent normal tissues were screened by the limma package and Venn diagram software. Functional assessment of DEGs was performed by Gene Ontology (GO) and the Kyoto Encyclopedia of Genes and Genomes (KEGG). Hub genes were selected by the protein-protein interaction (PPI) network and further validated by GSE14520 clinical data.

Results: A total of 26 up-regulated genes and 76 down-regulated genes were identified by analyzing three databases. GO and KEGG analysis demonstrated that these genes were involved in cell division, metabolism-related biological processes, the p53 pathway, and the cell cycle, among others. PPI network suggested that 14 hub DEGs (*TOP2A*, *HMMR*, *DTL*, *CCNB1*, *NEK2*, *PBK*, *RACGAP1*, *PRC1*, *CDK1*, *RRM2*, *ECT2*, *BUB1B*, *ANLN*, and *ASPM*) were most dysregulated and had potential to distinguish between HBV-associated HCC and noncancerous tissues. Further survival analysis of hub genes demonstrated that high expression of *TOP2A* was significantly associated with poor clinical outcomes of HBV-associated HCC.

Conclusions: *TOP2A* might serve as a key gene for prognosis and as a therapeutic target for HBV-associated HCC.

Keywords: Hepatitis B virus (HBV); hepatocellular carcinoma (HCC); differentially expressed gene (DEG); hub gene; *TOP2A*; Gene Ontology (GO); Kyoto Encyclopedia of Genes and Genomes (KEGG); protein-protein interaction (PPI); survival analysis; prognosis

Submitted Apr 02, 2021. Accepted for publication May 28, 2021.

doi: 10.21037/atm-21-2085

View this article at: https://dx.doi.org/10.21037/atm-21-2085

Introduction

Hepatocellular carcinoma (HCC) is one of the most common primary liver cancers, accounting for the third highest number of cancer-associated deaths worldwide (1,2). Multiple factors including hepatitis, diabetes, smoking, and alcohol consumption are known as risk factors of HCC (3,4). Among them, hepatitis B virus (HBV) infection serves as the leading cause contributing to the development and progression of HCC, which has been found to be related to 66% of cases (5). Although there have been great advances in HCC diagnosis and surgical techniques, patients with HBV-associated HCC have poor clinical prognosis due to virus induced-genetic alterations and irreversible hepatic

damage and cirrhosis (6,7). Therefore, further identification of genomic alterations of HBV-associated HCC is essential to provide potential targets for early diagnosis, as well as to develop novel therapeutic strategies.

Over the past decades, gene profiling and signatures, which can quickly detect differentially expressed genes (DEGs), have greatly accelerated cancer research. Several studies have analyzed the prognostic effects of array-based genes from HCC tumors. With genome-wide expression profiling, Cai *et al.* (8) developed a signature consisting of 11 genes that could effectively predict the overall survival (OS) of postoperative HCC patients. In addition, a 7-miRNAs-based signature was found to be significantly associated with recurrence-free survival in HCC (9). However, few investigations have identified gene signatures that predict poor prognosis for HBV-associated HCC. Thus, with public massive data and integrated bioinformatics methods, identifying important genes to predict prognosis in HBV-associated HCC is necessary and of great clinical significance (10).

With increasing use of high-throughput techniques, Gene Expression Omnibus (GEO) provides a public gene expression platform that contains millions of datasets and samples. It facilitates gene analysis including biomarker discovery, disease classification and phenotype comparisons. In the present study, we first reviewed HCC datasets in Gene Expression Omnibus (GEO), and 3 eligible datasets including patients with HBV-associated HCC were selected. Through overlapping analysis, a gene set with DEGs was identified. Of these DEGs, 14 were identified as hub DEGs in the protein-protein interaction (PPI) network. Different from previous studies identifying key biomarkers lacking of validation (11), we conducted validation analysis using GSE14520 clinical data which included HBV-associated tumor samples with complete prognostic information. The survival analysis of hub DEGs demonstrated that only high expression of *TOP2A* was significantly associated with both poor OS and disease-free survival (DFS). Therefore, *TOP2A* may serve as a biomarker for prognosis assessment and as a therapeutic target for HBV-associated HCC.

We present the following article in accordance with the REMARK reporting checklist (available at https://dx.doi.org/10.21037/atm-21-2085).

Methods

Gene datasets

The NCBI-GEO databases were searched to identify datasets that determined the DEGs in HCC. Eligible gene expression profiles of GSE 121248, GSE 62232, and GSE 55092 containing HBV-associated HCC and adjacent normal liver tissues were obtained. Microarray data of GSE121248, GSE62232, and GSE55092 were all based on GPL570 Platforms [(HG-U133_Plus_2) Affymetrix Human Genome U133 Plus 2.0 Array] which included 70 HBV-associated HCC tissues and 37 normal tissues, 10 HBV-associated HCC tissues and 10 normal tissues, and 39 HBV-associated HCC tissues and 81 normal tissues, respectively. All procedures were in accordance with the ethical standards of institutional and national committee on human experimentation and with the Helsinki Declaration (as revised in 2013).

Identification of DEGs

The genome expression profile was compared in the R platform using the limma package (12). Genes with $|\log FC| > 2$ and adjusted P value <0.05 between HBV-associated HCC and normal liver tissues were identified as DEGs. Volcano plots were constructed with the gplots package and overlapping genes were obtained via the online Venn diagram analysis.

Analysis of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways

GO is a community-based bioinformatics resource that supplies information regarding unique gene properties (13). KEGG is an integrated database resource for the biological interpretation of genomes, compounds, enzymes, diseases, drugs, and biological pathways. The online bioinformatics tool DAVID (14) was used to analyze GO enrichment and KEGG pathways (P<0.05).

Construction of the PPI network

For DEGs in the 3 cohorts, PPI networks for these genes were constructed with the online tools STRING and Cytoscape software (15). Moreover, the MCODE app in Cytoscape was used to assess the node distribution, path length distribution, and average clustering distribution. The top nodes with the highest degree of connectivity were identified as hub genes.

Survival analysis and identification of hub genes

In order to validate the DEGs that were associated with the

prognosis of HCC patients, hub genes in HBV-associated HCC were further assessed using GSE14520 clinical data. The OS and DFS of each hub gene were determined by Kaplan-Meier analysis. Finally, the correlation between clinicopathological parameters and significant hub genes was evaluated.

Statistical analysis

The test used to compare the expression between two groups was assessed by independent sample *t*-test. The FDR in DEG screening and GSEA were performed according to the Benjamini-Hochberg procedure. Volcano plots were constructed with the gplots package and overlapping genes were obtained via the online Venn diagram analysis. All statistical analyses were performed using SPSS, version 24.0 software with $P<0.05$ defined as statistical significance.

Results

Identification of DEGs in HBV-associated HCC

A total of 119 HBV-associated HCC tissues and 128 adjacent normal tissues from 3 datasets were included in our study. The whole-genome expression profile was compared in the R platform by using the limma package. A total of 145, 515, and 411 DEGs from GSE 121248, GSE 62232, and GSE 55092 were detected, respectively. The volcano plots of the DEGs are shown in *Figure 1A*,*B*,*C*. A Venn diagram was then used to obtain the overlapping genes within DEGs (*Figure 1D*,*E*). Finally, 102 overlapping genes were identified in HBV-associated HCC, including 26 up-regulated genes and 76 down-regulated genes (*Table 1*).

Functional evaluation of DEGs

The online bioinformatics tool DAVID software was then used to investigate 102 DEGs in terms of GO enrichment and KEGG pathways. The GO analysis demonstrated that up-regulated DEGs were significantly enriched in regulation of cell division, M phase, nuclear division, and mitosis for biological processes (BP), while down-regulated DEGs were enriched in oxidation reduction, secondary metabolic process, innate immune response, vitamin metabolic process, and immune response. For molecular function (MF), centrosome, microtubule organizing center, cytosol, and protein serine played the major roles in up-regulated DEGs, while electron carrier activity, iron ion, heme, tetrapyrrole, sugar, and carbohydrate binding were involved in down-regulated DEGs. In addition, spindle, cytoskeleton, and intracellular non-membrane-bounded organelle were the most significant cell components (CC) in up-regulated DEGs. Down-regulated DEGs were enriched in extracellular region, cell fraction, intrinsic and integral to plasma membrane, and endomembrane system (*Figure 2A*,*B* and Tables S1,S2).

In line with GO term enrichment results, KEGG analysis showed that DEGs were mainly responsible for metabolism-related biological processes including retinol, drug, caffeine, cytochrome P450, and tryptophan metabolism. P53 pathway and cytokine-cytokine receptor interaction were also involved in the development of HBV-associated HCC by DEGs (*Figure 2C* and *Table 2*).

Construction of the PPI network

To analyze the PPI, the selected 102 DEGs were covered. With the Cytoscape software, a PPI network was constructed with 70 nodes and 175 edges (*Figure 3A*). The top-ranked linker nodes included *TOP2A*, *HMMR*, *DTL*, *CCNB1*, *NEK2*, *PBK*, *RACGAP1*, *PRC1*, *CDK1*, *RRM2*, *ECT2*, *BUB1B*, *ANLN*, and *ASPM* via the MCODE app (*Figure 3B*). A simplified PPI network with the 14 up-regulated genes was generated. The selected hub genes in the cluster may have prognostic value for HBV-associated HCC.

Validation of hub genes

To confirm the effect of genes on survival, the 14 hub genes were further evaluated using GSE14520 clinical data. A total of 212 HBV-associated tumor samples and 210 adjacent normal samples with complete prognostic information were included. High expression of *ASPM*, *BUB1B*, *CDK1*, *NEK2*, *PBK*, *PRC1*, *RACGAP1*, *RPM2*, *TOP2A*, *ANLN*, and *DTL* were significantly associated with an increased risk of death (Figure S1). However, only patients with overexpression of *TOP2A* showed unfavorable recurrence-free survival (*Figure 4*). Further analysis based on age, gender, AFP level, tumor size, grade, and numbers in the GSE14520 cohort indicated that high expression of *TOP2A* was associated with an increased level of AFP (*Figure 5* and *Table 3*).

Discussion

According to the estimates of recent study, about 350

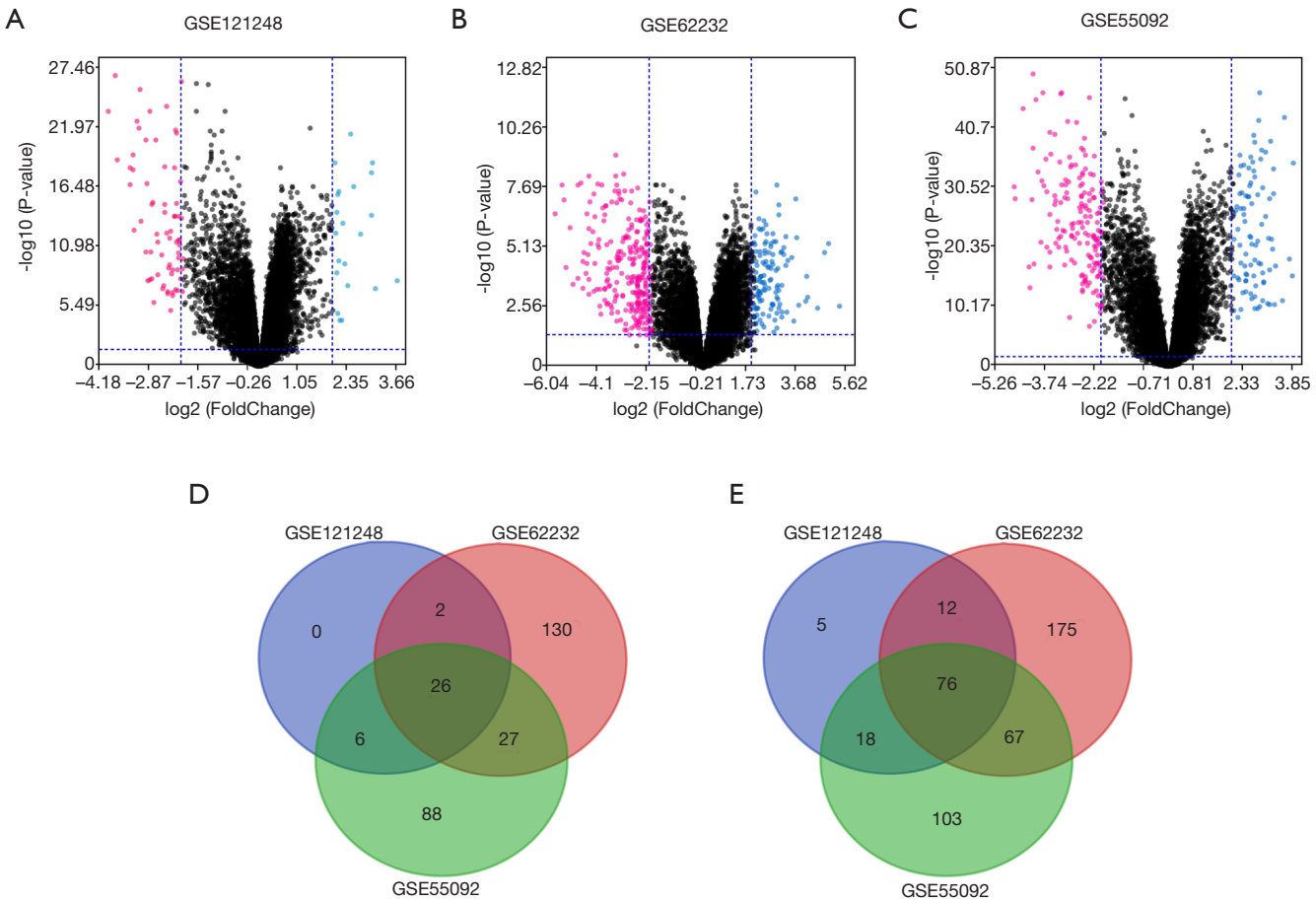

Figure 1 Demonstration of the DEGs in the 3 databases (GSE121248, GSE62232, and GSE55092). (A,B,C) Volcano plot of DEGs. Blue dots: up-regulation; red dots: down-regulation; black dots: non-differentially expressed genes. (D) A total of 26 DEGs were up-regulated and (E) 76 DEGs were down-regulated. DEGs, differentially expressed genes.

Table 1 A total of 102 DEGs were detected from 3 databases, including 26 up-regulated genes and 76 down-regulated genes in the HBV-associated HCC tissues compared to normal tissues

DEGs	Gene names
Up-regulated	*SPINK1, CAP2, DTL, IGF2BP3, CCNB1, ASPM, HMMR, AKR1B10, GPC3, ROBO1, SPP1, ZIC2, ANLN, COL15A1, PRC1, CDK1, RACGAP1, RRM2, TOP2A, PBK, SULT1C2, NEK2, ACSL4, CRNDE, BUB1B, ECT2*
Down-regulated	*CYP4A22, CYP26A1, BBOX1, CYP2A6, CNTN3, TENM1, LINC01093, CXCL14, SLC22A1, IGF1, CYP39A1, HAO2, FAM134B, MT1F, SLC25A47, MFSD2A, ZG16, HHIP, KCNN2, SLCO1B3, CYP1A2, CNDP1, BCO2, FCN3, GBA3, TTC36, CLEC4G, C3P1, CYP2B6, GYS2, KMO, CD5L, LPA, GHR, CLEC1B, CXCL2, MIR675, FOSB, LIFR, FAM65C, CYP2C9, CLRN3, LCAT, CLEC4M, VNN1, ESR1, PLAC8, ALDOB, HAMP, DNASE1L3, DCN, NAT2, IL1RAP, AKR1D1, CXCL12, TMEM27, CRHBP, THRSP, IDO2, HGFAC, ADGRG7, C7, FREM2, ADH4, GPM6A, OIT3, MT1M, HGF, GLYAT, CYP2B7P, GLS2, ADRA1A, APOF, C9, SRPX, FCN2*

DEGs, differentially expressed genes; HBV, hepatitis B virus; HCC, hepatocellular carcinoma.

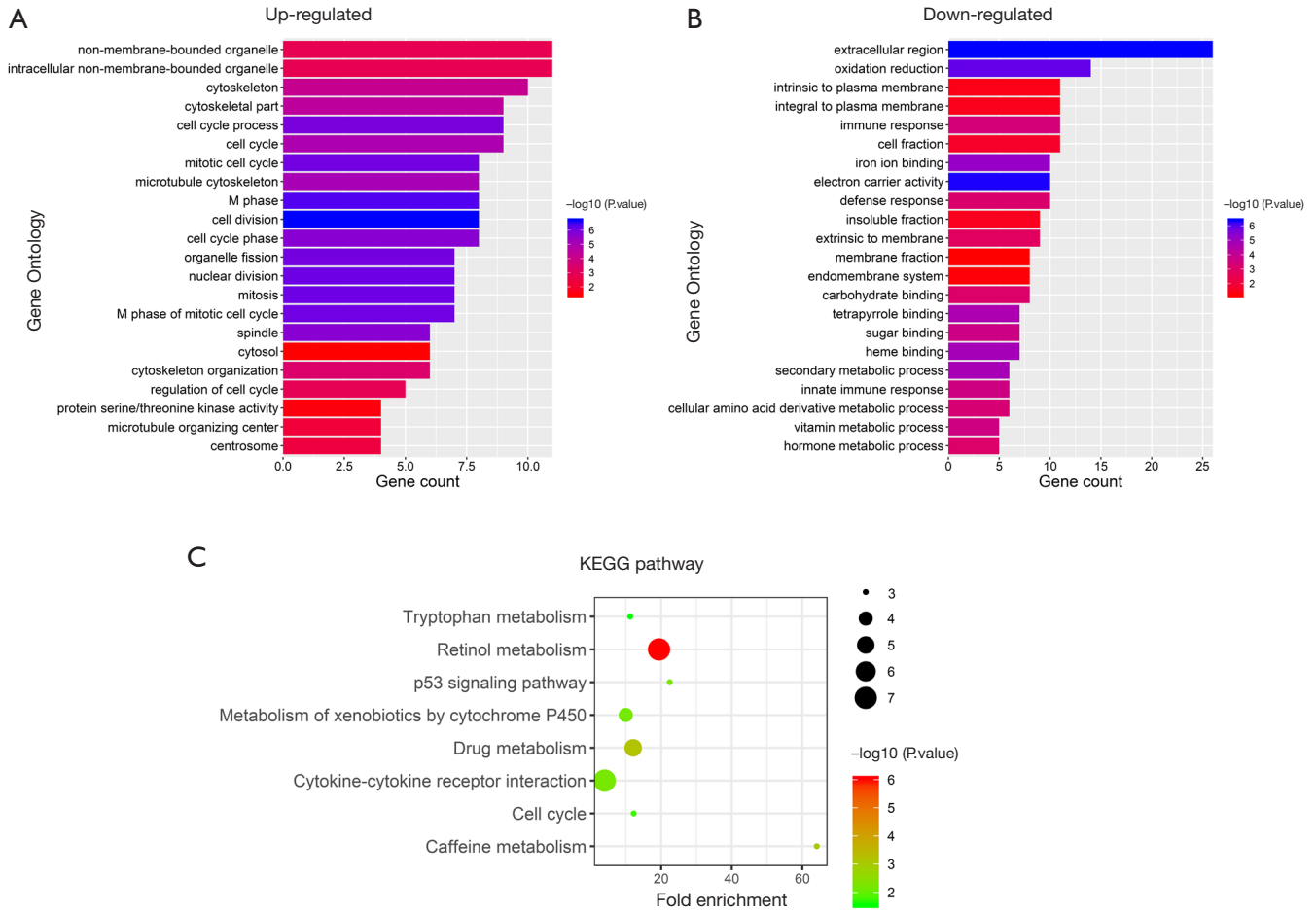

Figure 2 GO and KEGG enrichment results of 102 DEGs. (A,B) GO term enrichment results of up-regulated DEGs and down-regulated DEGs; (C) KEGG results of DEGs. GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEGs, differentially expressed genes.

Table 2 KEGG pathway analysis of differentially expressed genes in HBV-associated HCC

Term	Genes
p53 signaling pathway	*CCNB1, CDK1, RRM2*
Cell cycle	*CCNB1, CDK1, BUB1B*
Retinol metabolism	*CYP4A22, -2C9, -2B6, -26A1, 2A6, 1A2, ADH4*
Drug metabolism	*CYP2C9, -2B6, -2A6, -1A2, ADH4*
Caffeine metabolism	*NAT2, CYP2A6, -1A2*
Cytokine-cytokine receptor interaction	*CXCL14, CXCL2, IL1RAP, LIFR, HGF, CXCL12, GHR*
Metabolism of xenobiotics by cytochrome P450	*CYP2C9, -2B6, -1A2, ADH4*
Tryptophan metabolism	*IDO2, KMO, CYP1A2*

KEGG, Kyoto Encyclopedia of Genes and Genomes; HBV, hepatitis B virus; HCC, hepatocellular carcinoma.

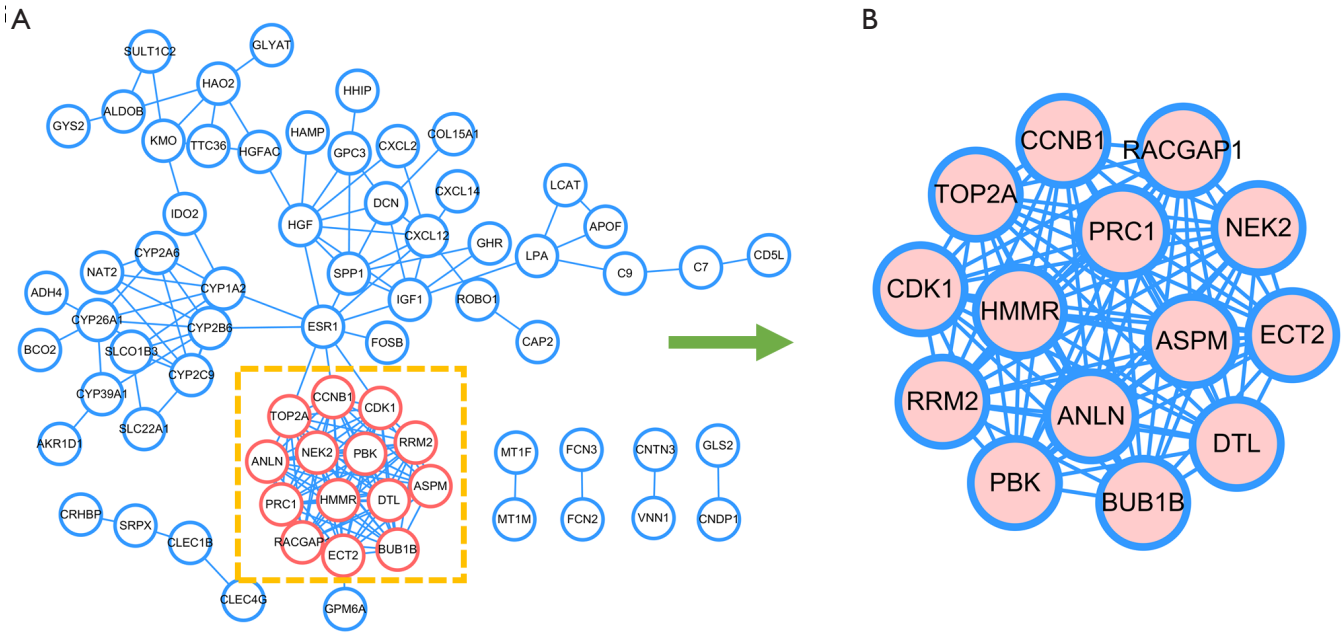

Figure 3 Construction of DEGs PPI network and identification of hub genes. (A) The PPI network between DEGs was constructed by the online STRING database; (B) a total of 14 hub genes were identified via Cytoscape software (degree cutoff =2, node score cutoff =0.2, k-core =2, and max. Depth =100). DEGs, differentially expressed genes; PPI, protein-protein interaction.

million people worldwide are infected with HBV. HBV infection, the leading etiology of HCC, accounts for 66% cases of HCC incidence globally (16). HBV infection is mainly transmitted vertically in endemic areas, especially in developing countries. Accordingly, the average age of HBV carriers who develop HCC is younger than that of other etiologies. Besides, liver inflammation caused by immune responses during HBV infection leads to liver fibrosis and cirrhosis in majority of patients, which promotes the development of HCC. In addition, integration of HBV DNA into the host genome induces both genetic instability and mutagenesis of various cancer-related genes, which is pathologically unique compared with other types of HCC (17,18). Since the key genes of HBV-associated HCC are different from other types of HCC, we conducted this study to explore DEGs for their novel prognostic value and clinical significance in HBV-associated HCC.

In the present study, we identified 26 up-regulated genes and 76 down-regulated genes which were involved in cell division, metabolism-related biological processes, the p53 pathway, and the cell cycle. Additionally, 14 hub DEGs were further validated and *TOP2A* was found to be associated with unfavorable OS and recurrence-free survival in HBV-associated cohorts. *TOP2A*, also known as DNA topoisomerase II alpha, is highly expressed at G2/M phase and a key regulator of DNA decatenation during mitosis (19,20). As a marker of proliferation in malignant cells, *TOP2A* has been found correlated with poor prognosis in various types of cancers such as prostate cancer, breast cancer, and gallbladder carcinoma (21-23). Our results demonstrated higher expression of *TOP2A* in HBV-associated HCC tissues and worse survival outcomes induced by *TOP2A*. The results were consistent with previous studies in HCC (24,25). Recent study by Kwan *et al.* (26) showed that *TOP2A* was a downstream target of *TRRAP* and *KAT* regulating HCC cell growth. Depletion of *TOP2A* caused reduced colony formation, induction of senescence and G2/M arrest of HCC, which confirmed the cancer promoting role of *TOP2A*. Panvichian *et al.* further reported that *TOP2A* in HCC was related to the presence of hepatitis B surface antigen in serum (27). This may explain why *TOP2A* was uniquely identified from hub genes in HBV-associated HCC. In addition, an elevated level of AFP was also found in patients with high *TOP2A* expression, which is in line with other malignant characteristics including microvascular invasion and chemotherapy resistance reported by other studies (28). Thus, high expression of *TOP2A* may serve as one of the

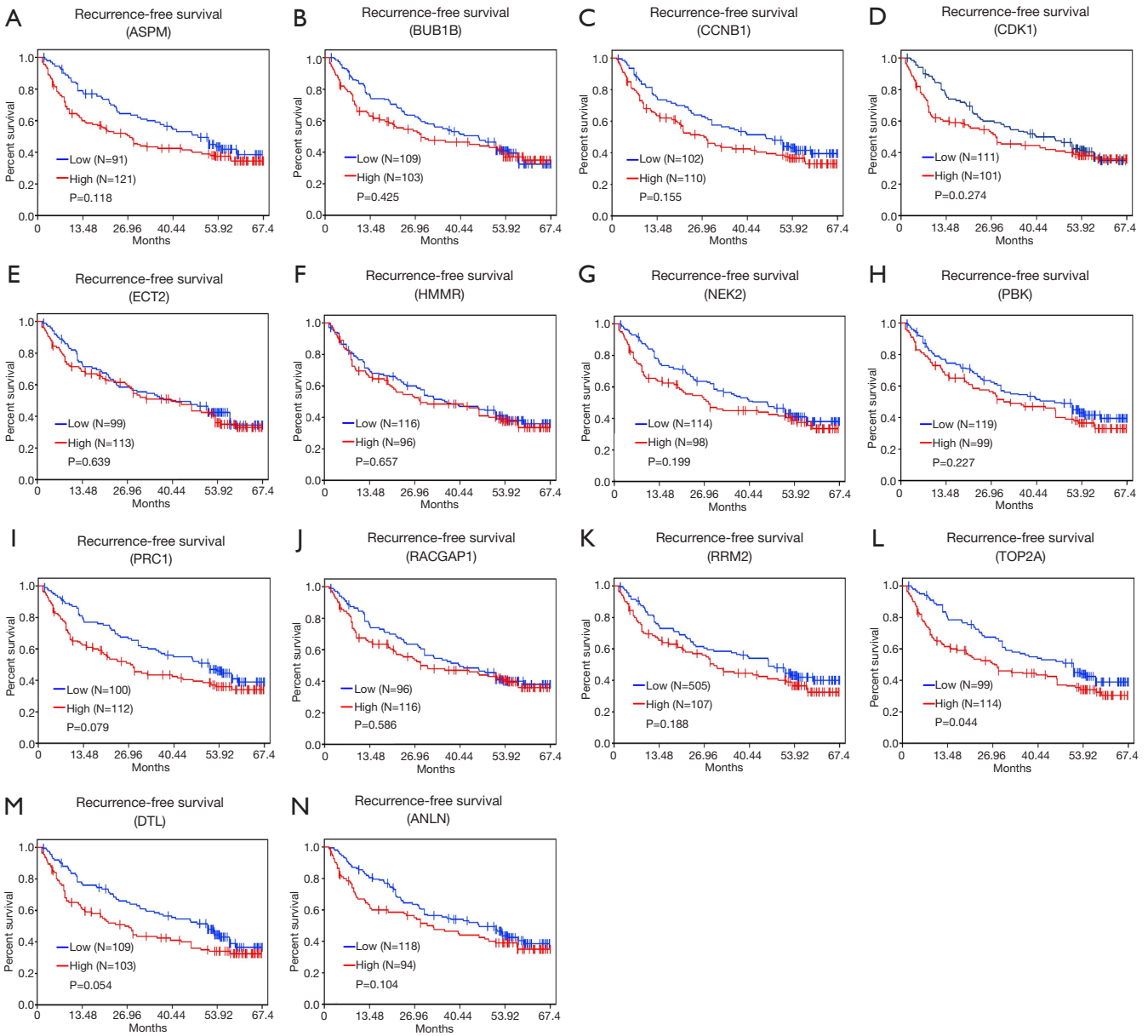

Figure 4 The recurrence-free survival analysis of 14 hub genes by GSE14520 clinical data. Only patients with overexpression of TOP2A showed unfavorable recurrence-free survival (P<0.05).

genomic hallmarks for prognosis and as a therapeutic target for HBV-associated HCC.

Recently, multiple studies have identified crucial genes for HCC. DEGs including *CCNB1*, *CDC20*, *PRC1*, *NDC80*, and *FOXM1*, among others, were found to be up-regulated in tumor tissues and predicted poor prognosis (29,30). The difference between our study and previous studies was that we focused on a genetically different type of HCC. Our current study selected gene profiles from HBV-associated tissues, while other studies included HCC samples which originated from diverse etiologies including hepatitis infection, autoimmune disease, and nonalcoholic steatohepatitis, which resulted in different gene signatures (31). In addition, another advantage in our study was that hub genes identified by the PPI network were further validated in the biggest online HBV-associated

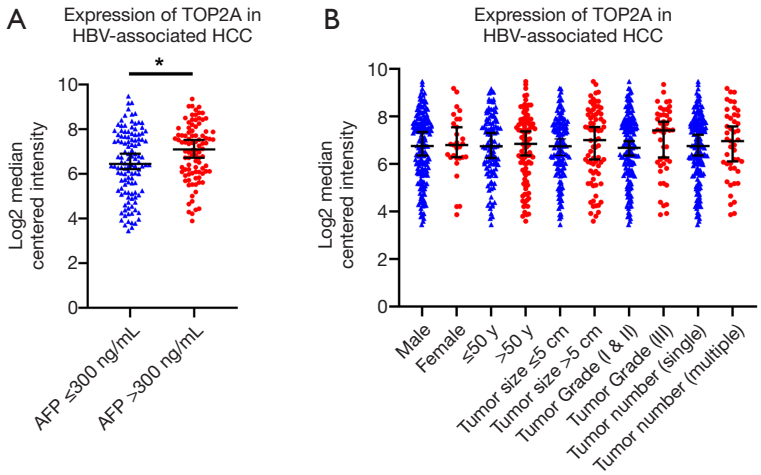

Figure 5 Correlation between clinicopathological characteristics and the expression of TOP2A in HBV-associated HCC patients of GSE14520. (A) Patients with an elevated level of AFP (>300 ng/mL) had significantly higher expression of TOP2A in HBV-associated HCC (*, P<0.05); (B) clinical parameters of age, gender, tumor size, grade, and numbers were not correlated with the expression of TOP2A in HBV-associated HCC. HBV, hepatitis B virus; HCC, hepatocellular carcinoma.

Table 3 Correlation between expression of TOP2A and clinicopathological parameters in HBV-associated HCC

Clinicopathological parameters	Log2 expression of TOP2A	P value
Gender		
Male	6.67±1.42	0.6920
Female	6.80±1.30	
Age (years)		
<50	6.68±1.40	0.9339
≥50	6.69±1.42	
Tumor size (cm)		
<5	6.64±1.31	0.3573
≥5	6.77±1.57	
AFP (ng/mL)		
<300	6.46±1.47	0.0143
≥300	6.98±1.26	
Tumor grade		
I & II	6.60±1.39	0.0933
III	6.97±1.44	
Tumor number		
Single	167	0.5359
Multiple	45	

HBV, hepatitis B virus; HCC, hepatocellular carcinoma.

cohort based on GSE14520. *TOP2A*, with significant prognostic value, was revealed by survival analysis, which is more reliable compared with previous studies.

Certain underlying limitations need to be considered when interpreting our results. Firstly, our study only enrolled 3 gene expression profiles from GEO datasets comprising 119 HCC tissues and 128 adjacent normal tissues. The 3 GEO datasets lack expressions of microRNA, lncRNA and circRNA as well. The limited availability of HBV-associated HCC datasets impeded us from accurately screening DEGs. Also, due to the absence of HBV infection status in TCGA datasets, our current study failed to validate hub genes in TCGA cohorts. In addition, clinicopathological parameters from GSE14520 including age, gender, tumor size, grade, and number were not sufficient. More tumor characteristics are required to evaluate the prognostic value of key genes. Besides, the basic experimental data validating the role of *TOP2A* is lacking. Further experiments both *in vivo* and *in vitro* are essential to confirm the prognostic value of *TOP2A* in HBV-associated HCC.

Conclusions

Despite underlying limitations, our present study identified 102 DEGs in HBV-associated HCC. GO and KEGG analysis demonstrated increased division and proliferation of HBV-associated cancer cells induced by these DEGs. Using the PPI network and further investigations, we found that overexpression of *TOP2A* was significantly associated with poor prognosis and increased AFP level. *TOP2A* might serve as a key gene for prognosis and as a therapeutic target for HBV-associated HCC.

Acknowledgments

Funding: This study was supported by National Key Research on Precision Medicine of China (2017YFC0908102, 2018ZX10723204) and National Natural Science Foundation of China (81902379).

Footnotes

Reporting Checklist: The authors have completed the REMARK reporting checklist. Available at https://dx.doi.org/10.21037/atm-21-2085

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at https://dx.doi.org/10.21037/atm-21-2085). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. All procedures were in accordance with the ethical standards of institutional and national committee on human experimentation and with the Helsinki Declaration (revised in 2013). There was no interaction with patients directly, as we acquired data from online public datasets.

References

1. Forner A, Reig M, Bruix J. Hepatocellular carcinoma. Lancet 2018;391:1301-14.
2. Bray F, Ferlay J, Soerjomataram I, et al. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018;68:394-424.
3. Wang M, Xi D, Ning Q. Virus-induced hepatocellular carcinoma with special emphasis on HBV. Hepatol Int 2017;11:171-80.
4. Fujiwara N, Friedman SL, Goossens N, et al. Risk factors and prevention of hepatocellular carcinoma in the era of precision medicine. J Hepatol 2018;68:526-49.
5. Stanaway JD, Flaxman AD, Naghavi M, et al. The global burden of viral hepatitis from 1990 to 2013: findings from the Global Burden of Disease Study 2013. Lancet 2016;388:1081-8.
6. Grandhi MS, Kim AK, Ronnekleiv-Kelly SM, et al. Hepatocellular carcinoma: From diagnosis to treatment. Surg Oncol 2016;25:74-85.
7. Genco C, Cabibbo G, Maida M, et al. Treatment of hepatocellular carcinoma: present and future. Expert Rev Anticancer Ther 2013;13:469-79.

8. Cai J, Li B, Zhu Y, et al. Prognostic Biomarker Identification Through Integrating the Gene Signatures of Hepatocellular Carcinoma Properties. EBioMedicine 2017;19:18-30.
9. Wang G, Dong F, Xu Z, et al. MicroRNA profile in HBV-induced infection and hepatocellular carcinoma. BMC Cancer 2017;17:805.
10. Roessler S, Long EL, Budhu A, et al. Integrative genomic identification of genes on 8p associated with hepatocellular carcinoma progression and patient survival. Gastroenterology 2012;142:957-966.e12.
11. Liao X, Yu T, Yang C, et al. Comprehensive investigation of key biomarkers and pathways in hepatitis B virus-related hepatocellular carcinoma. J Cancer 2019;10:5689-704.
12. Ritchie ME, Phipson B, Wu D, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. Nucleic Acids Res 2015;43:e47.
13. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat Genet 2000;25:25-9.
14. Huang W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. Nat Protoc 2009;4:44-57.
15. Shannon P, Markiel A, Ozier O, et al. Cytoscape: a software environment for integrated models of biomolecular interaction networks. Genome Res 2003;13:2498-504.
16. Kulik L, El-Serag HB. Epidemiology and Management of Hepatocellular Carcinoma. Gastroenterology 2019;156:477-491.e1.
17. Yang JD, Hainaut P, Gores GJ, et al. A global view of hepatocellular carcinoma: trends, risk, prevention and management. Nat Rev Gastroenterol Hepatol 2019;16:589-604.
18. Schweitzer A, Horn J, Mikolajczyk RT, et al. Estimations of worldwide prevalence of chronic hepatitis B virus infection: a systematic review of data published between 1965 and 2013. Lancet 2015;386:1546-55.
19. D Arcy N, Gabrielli B. Topoisomerase II Inhibitors and Poisons, and the Influence of Cell Cycle Checkpoints. Curr Med Chem 2017;24:1504-19.
20. Delgado JL, Hsieh CM, Chan NL, et al. Topoisomerases as anticancer targets. Biochem J 2018;475:373-98.
21. Labbé DP, Sweeney CJ, Brown M, et al. TOP2A and EZH2 Provide Early Detection of an Aggressive Prostate Cancer Subgroup. Clin Cancer Res 2017;23:7072-83.
22. Washiro M, Ohtsuka M, Kimura F, et al. Upregulation of topoisomerase IIalpha expression in advanced gallbladder carcinoma: a potential chemotherapeutic target. J Cancer Res Clin Oncol 2008;134:793-801.
23. Romero A, Martín M, Cheang MC, et al. Assessment of Topoisomerase II α status in breast cancer by quantitative PCR, gene expression microarrays, immunohistochemistry, and fluorescence in situ hybridization. Am J Pathol 2011;178:1453-60.
24. Zhang L, Huang Y, Ling J, et al. Screening and function analysis of hub genes and pathways in hepatocellular carcinoma via bioinformatics approaches. Cancer Biomark 2018;22:511-21.
25. Wu M, Liu Z, Zhang A, et al. Identification of key genes and pathways in hepatocellular carcinoma: A preliminary bioinformatics analysis. Medicine (Baltimore) 2019;98:e14287.
26. Kwan SY, Sheel A, Song CQ, et al. Depletion of TRRAP Induces p53-Independent Senescence in Liver Cancer by Down-Regulating Mitotic Genes. Hepatology 2020;71:275-90.
27. Panvichian R, Tantiwetrueangdet A, Angkathunyakul N, et al. TOP2A amplification and overexpression in hepatocellular carcinoma tissues. Biomed Res Int 2015;2015:381602.
28. Wong N, Yeo W, Wong WL, et al. TOP2A overexpression in hepatocellular carcinoma correlates with early age onset, shorter patients survival and chemoresistance. Int J Cancer 2009;124:644-52.
29. Shen S, Kong J, Qiu Y, et al. Identification of core genes and outcomes in hepatocellular carcinoma by bioinformatics analysis. J Cell Biochem 2019;120:10069-81.
30. Wang J, Tian Y, Chen H, et al. Key signaling pathways, genes and transcription factors associated with hepatocellular carcinoma. Mol Med Rep 2018;17:8153-60.
31. Zhou L, Du Y, Kong L, et al. Identification of molecular target genes and key pathways in hepatocellular carcinoma by bioinformatics analysis. Onco Targets Ther 2018;11:1861-9.

(English Language Editor: C. Betlazar-Maseh)

Cite this article as: Sha M, Cao J, Zong ZP, Xu N, Zhang JJ, Tong Y, Xia Q. Identification of genes predicting unfavorable prognosis in hepatitis B virus-associated hepatocellular carcinoma. Ann Transl Med 2021;9(12):975. doi: 10.21037/atm-21-2085

Supplementary

Table S1 Gene ontology analysis of up-regulated expressed genes in HBV-associated HCC

Category	Term	Count	P value	FDR
GOTERM_BP_DIRECT	Cell division	8	1.95E-07	2.71E-04
	M phase	8	4.09E-07	5.68E-04
	Mitosis	7	7.65E-07	0.00106
	Nuclear division	7	7.65E-07	0.00106
	M phase of mitotic cycle	7	8.50E-07	0.00118
	Mitotic cell cycle	8	9.04E-07	0.00125
	Organelle fission	7	9.67E-07	0.00134
	Cell cycle process	9	1.11E-06	0.00154
	Cell cycle	9	1.18E-05	0.01634
	Cytoskeleton organization	6	4.51E-04	0.62344
	Regulation of cell cycle	5	0.00151	2.08031
GOTERM_CC_DIRECT	Spindle	6	1.92E-06	0.00213
	Microtubule cytoskeleton	8	8.33E-06	0.00925
	Cytoskeleton	10	6.58E-05	0.073050
	Intracellular non-membrane-bounded organelle	11	0.00177	1.94895
	Non-membrane-bounded organelle	11	0.00177	1.94895
GOTERM_MF_DIRECT	Centrosome	4	0.00418	4.54495
	Microtubule organizing center	4	0.00587	6.32857
	Cytosol	6	0.0407	36.97028
	Protein serine/ threonine kinase activity	4	0.0308	29.05918

https://dx.doi.org/10.21037/atm-21-2085

Table S2 Gene ontology analysis of down-regulated expressed genes in HBV-associated HCC

Category	Term	Count	P value	FDR
GOTERM_BP_DIRECT	Oxidation reduction	14	1.20E-06	0.00182
	Secondary metabolic process	6	1.65E-05	0.02527
	Innate immune response	6	2.37E-04	0.36189
	Vitamin metabolic process	5	2.40E-04	0.36594
	Immune response	11	4.06E-04	0.61801
	Cellular amino acid derivative metabolic process	6	5.55E-04	0.84488
	Defense response	10	7.63E-04	1.15976
	Hormone metabolic process	5	8.93E-04	1.35635
GOTERM_CC_DIRECT	Extracellular region	26	4.10E-07	4.64E-04
	Extrinsic to membrane	9	0.00147	1.64552
	Cell fraction	11	0.02026	20.65319
	Insoluble fraction	9	0.03165	30.46864
	Integral to plasma membrane	11	0.03566	33.65317
	Intrinsic to plasma membrane	11	0.04069	37.45891
	Endomembrane system	8	0.05785	48.99586
	Membrane fraction	8	0.06697	54.30599
GOTERM_MF_DIRECT	Electron carrier activity	10	4.51E-07	5.71E-04
	Iron ion binding	10	7.00E-06	0.00886
	Heme binding	7	1.42E-05	0.01801
	Tetrapyrrole binding	7	2.05E-05	0.02594
	Sugar binding	7	2.04E-04	0.25768
	Carbohydrate binding	8	8.64E-04	1.08720

https://dx.doi.org/10.21037/atm-21-2085

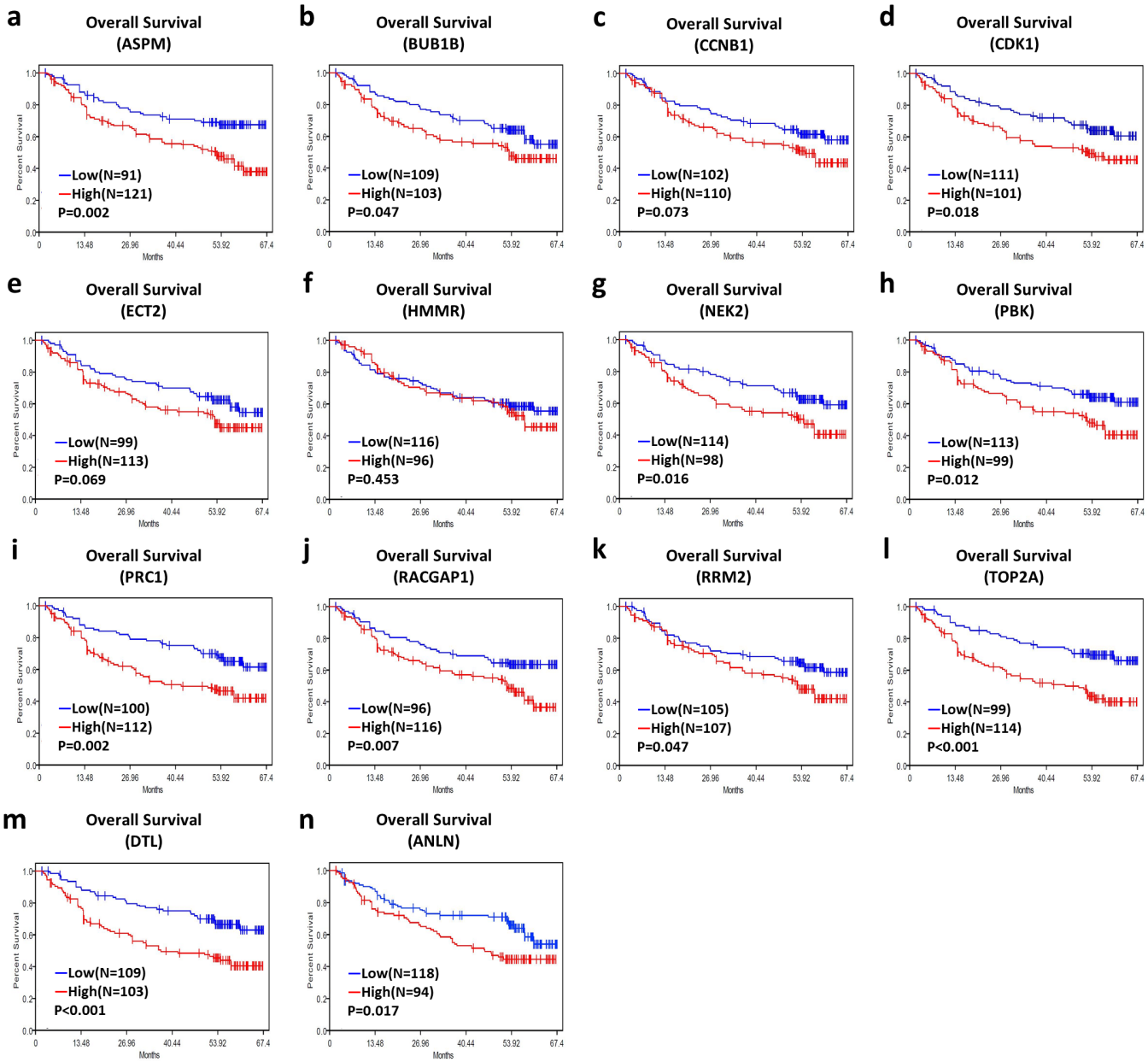

Figure S1 The overall survival analysis of 14 hub genes by GSE14520 clinical data. Overexpression of 11 genes (*ASPM*, *BUB1B*, *CDK1*, *NEK2*, *PBK*, *PRC1*, *RACGAP1*, *RPM2*, *TOP2A*, *ANLN* and *DTL*) had a significantly worse survival rate (P<0.05).

https://dx.doi.org/10.21037/atm-21-2085