

# Exploring new targets for the treatment of hepatitis-B virus and hepatitis-B virus-associated hepatocellular carcinoma

## A new perspective in bioinformatics

Yang Wang, PhD<sup>1</sup>, ShanShan Wang, PhD, Yang Che, MM, DeXi Chen, PhD, YaLi Liu, PhD, Ying Shi, PhD\*

### Abstract

**Background:** Hepatitis B Virus (HBV) infection is a global public health problem. After infection, patients experience a natural course from chronic hepatitis to cirrhosis and even Hepatitis B associated Hepatocellular Carcinoma (HBV-HCC). With the multi-omics research, many differentially expressed genes from chronic hepatitis to HCC stages have been discovered. All these provide important clues for new biomarkers and therapeutic targets. The purpose of this study is to explore the differential gene expression of HBV and HBV-related liver cancer, and analyze their enrichments and significance of related pathways.

**Methods:** In this study, we downloaded four microarray datasets GSE121248, GSE67764, GSE55092, GSE55092 and GSE83148 from the Gene Expression Omnibus (GEO) database. Using these four datasets, patients with chronic hepatitis B (CHB) differentially expressed genes (CHB DEGs) and patients with HBV-related HCC differentially expressed genes (HBV-HCC DEGs) were identified. Then Protein-protein Interaction (PPI) network analysis, Gene Ontology (GO) functional analysis and Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis were performed to excavate the functional interaction of these two groups of DEGs and the common DEGs. Finally, the Kaplan website was used to analyze the role of these genes in HCC prognostic.

**Results:** A total of 241 CHB DEGs, 276 HBV-HCC DEGs, and 4 common DEGs (cytochrome P450 family 26 subfamily A member 1 (CYP26A1), family with sequence similarity 110 member C (FAM110C), SET and MYND domain containing 3 (SMYD3) and zymogen granule protein 16 (ZG16)) were identified. CYP26A1, FAM110C, SMYD3 and ZG16 exist in 4 models and interact with 33 genes in the PPI network of CHB and HBV-HCC DEGs. GO function analysis showed that: CYP26A1, FAM110C, SMYD3, ZG16, and the 33 genes in their models mainly affect the regulation of synaptic vesicle transport, tangential migration from the subventricular zone to the olfactory bulb, cellular response to manganese ion, protein localization to mitochondrion, cellular response to dopamine, negative regulation of neuron death in the biological process of CHB. In the biological process of HBV-HCC, they mainly affect tryptophan catabolic process, ethanol oxidation, drug metabolic process, tryptophan catabolic process to kynurenine, xenobiotic metabolic process, retinoic acid metabolic process, steroid metabolic process, retinoid metabolic process, steroid catabolic process, retinal metabolic process, and rogen metabolic process. The analysis of the 4 common DEGs related to the prognosis of liver cancer showed that: CYP26A1, FAM110C, SMYD3 and ZG16 are closely related to the development of liver cancer and patient survival. Besides, further investigation of the research status of the four genes showed that CYP26A1 and SMYD3 could also affect HBV replication and the prognosis of liver cancer.

**Conclusion:** CYP26A1, FAM110C, SMYD3 and ZG16 are unique genes to differentiate HBV infection and HBV-related HCC, and expected to be novel targets for HBV-related HCC occurrence and prognostic judgement.

**Abbreviations:** CHB DEGs = chronic hepatitis B differentially expressed genes, CYP26A1 = cytochrome P450 family 26 subfamily A member 1, DAVID = Database for Annotation Visualization and Integrated Discovery, FAM110C = family with sequence similarity 110 member C, FC = foldchange, GEO = Gene Expression Omnibus, GO = Gene Ontology, HBV = hepatitis B, HBV-HCC = hepatitis B associated hepatocellular carcinoma, HBV-HCC = hepatitis B associated hepatocellular carcinoma, HBV-HCC DEGs =

Editor: Ning Zhang.

Research and demonstration application of clinical diagnosis and treatment technology in the capital (Z191100006619064); Beijing Natural Science Foundation (7182072).

The manuscript is an original study presenting novel work, and it has not been previously submitted or accepted by any other journal. The manuscript has been approved by all authors, and obtained ethics approval and written informed consent. The authors who take part in this study declare that they do not have anything to disclose regarding funding or conflicts of interest with respect to this manuscript.

The datasets generated during and/or analyzed during the current study are publicly available.

Beijing Institute of Hepatology, Beijing You'an Hospital, Capital Medical University, Beijing, China.

\* Correspondence: Ying Shi, Beijing Institute of Hepatology, Beijing You'an Hospital, Capital Medical University, Beijing, 100069, China (e-mail: yingshi@ccmu.edu.cn).

Copyright © 2021 the Author(s). Published by Wolters Kluwer Health, Inc.

This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Wang Y, Wang S, Che Y, Chen D, Liu Y, Shi Y. Exploring new targets for the treatment of hepatitis-B virus and hepatitis-B virus-associated hepatocellular carcinoma: a new perspective in bioinformatics. *Medicine* 2021;100:33(e26917).

Received: 2 March 2021 / Received in final form: 30 June 2021 / Accepted: 24 July 2021

<http://dx.doi.org/10.1097/MD.00000000000026917>

hepatitis B associated hepatocellular carcinoma differentially expressed genes, HCC = hepatocellular carcinoma, KEGG = Kyoto Encyclopedia of Genes and Genomes, PPI = protein–protein interaction network, SMYD3 = SET and MYND domain containing 3, STRING = Search Tool for the Retrieval of Interacting Genes, ZG16 = Zymogen granule protein 16.

**Keywords:** bioinformatics, Cytochrome P450 family 26 subfamily A member 1, Family with sequence similarity 110 member C, hepatitis B virus, hepatitis B virus-related hepatocellular carcinoma, hepatocellular carcinoma, SET and MYND domain containing 3, Zymogen granule protein 16

## 1. Introduction

Hepatitis B virus (HBV) virus is a hepatotropic DNA virus that is the main cause of high morbidity and mortality of liver-related diseases worldwide. Currently, >2 billion people have been exposed to HBV, accounting for 30% of all deaths from cirrhosis and 40% of all deaths related to hepatocellular carcinoma globally.<sup>[1,2]</sup> Hepatocellular carcinoma is the fifth most common type of malignancy and the most prevalent cancers worldwide.<sup>[2]</sup> Its most common etiological factor world is HBV infection worldwide, especially in China.<sup>[3]</sup> Although the current oral antiviral drugs can suppress the viral load and reduce liver-related complications, lifetime treatment, high cost, and potential toxicity are still major issues that need to be faced in the treatment process. Therefore, it is necessary to develop new treatment approaches to tackle the shortcoming of HBV treatment.<sup>[4]</sup> With the application of multi-omics analysis methods, numerous important molecules have been discovered and successfully applied clinically for diagnosis and treatment. Since HBV infection is the major cause of liver cancer worldwide, it must be considered as an important factor in the study of liver cancer.

It is a very long process from HBV infection to the development of HBV-HCC, and the mechanism is very complicated and profound. The research in this field is diverse and extensive, especially in bioinformatics data. In our research, we focus on the gene changes between HBV and HBV-HCC, to explore the similarities and differences between the 2 pathological processes. Through this study, we hope to find some new genes as targets for the diagnosis and prognosis of HBV-HCC.

## 2. Material and methods

### 2.1. Microarray data

The GSE121248, GSE67764, GSE55092, GSE47197, and GSE83148 gene expression datasets (Table 1) were downloaded from the Gene expression Omnibus (GEO) database (<http://www.ncbi.nlm.nih.gov/geo>). The datasets generated during and/or analyzed during the study are available in the GEO database (<http://www.ncbi.nlm.nih.gov/geo>). Based on the annotation information in the platform, the probes were converted into

the corresponding gene symbol. The platform of the GSE121248 and GSE55092 are the GPL570 [HG-U133\_Plus\_2] Affymetrix Human Genome U133 Plus 2.0 Array. The GSE121248 contains 70 chronic hepatitis B induced HCC tissues and 37 CHB tissues. The GSE55092 contains 11 HBV-associated HCC tissues and paired 11 CHB tissues. The platform of the GSE67764 dataset is the GPL17077 Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray, and this dataset contains 3 normal liver tissues, 3 chronic hepatitis B induced HCC tissues, and 3 CHB tissues. The GSE47197, platform is GPL16699 Agilent-039494 SurePrint G3 Human GE v2 8x60K Microarray, containing 62 chronic hepatitis B induced HCC tissues and paired 62 CHB tissues. The platform of the GSE83148 dataset is the Affymetrix Human Genome U133 Plus 2.0 Array, containing 116 HBV infected samples and 6 normal control samples (Table 1).

### 2.2. Identification of DEGs

GEO2R (<http://www.ncbi.nlm.nih.gov/geo/geo2r>) was used to screen DEGs in these 4 datasets. The *P* values were adjusted to correct false-positive results using the Benjamini and Hochberg False Discovery Rate method. Genes with more than one probe set or probe sets without relevant gene symbols were removed or averaged, respectively. Genes with a *adj.P*-value <.01 and  $|\log_{2}FC$  (foldchange)|>1 were considered DEGs.

### 2.3. GO and KEGG pathway enrichment analyses of DEGs

The Database for Annotation, Visualization and Integrated Discovery (DAVID; <http://david.ncifcrf.gov>)<sup>[5]</sup> that provides a comprehensive set of functional annotation information of genes and proteins, was used to analyze the DEGs enrichment and functional annotation. GO annotation was performed to analyze the biological process of DEGs.<sup>[6]</sup> Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis to understand high-level functions and biological systems from large-scale molecular datasets generated by high-throughput experiments.<sup>[7]</sup> The *P*<.05 was considered statistically significant.

### 2.4. PPI network construction and module analysis

To search the interactions and functions in the DEGs, the Search Tool for the Retrieval of Interacting Genes (STRING; <http://string-db.org>) was used to analyze the functional associations and construct a PPI network. In this study, an interaction with a combined score >0.4 was considered statistically significant. The PPI network was constructed by Cytoscape (version 3.7.1), is a source bioinformatics software platform for visualizing molecular interaction networks.<sup>[8]</sup> The Molecular Complex Detection (MCODE) plugin in Cytoscape software (the U.S. National Institute of General Medical Sciences (NIGMS), America)<sup>[9]</sup> was applied to select the significant modules of hub genes from the PPI network (degree cut-off ≥ 2, node score cut-off ≥ 0.1, max depth

**Table 1**

The number of specimens in each dataset.

Dataset	Normal liver tissues	CHB tissues	HBV-HCC tissues
GSE121248	0	37	70
GSE55092	0	11	11
GSE67764	3	3	3
GSE47197	0	62	62
GSE83148	6	116	0

CHB = chronic hepatitis B; HBV-HCC = hepatitis B-associated hepatocellular carcinoma.

D 100, and K-core  $\geq 2$ ). The KEGG pathway and GO analyses for DEGs in modules were used to investigate their potential information using DAVID.

### 2.5. Key genes selection, cluster analysis, and survival analysis

The plug-in Biological Networks Gene Oncology tool (version 3.0.3)<sup>[10]</sup> in Cytoscape software was used to illustrate the biological process analysis of key genes and the significant model. In addition, the Kaplan website (<https://kmplot.com/analysis>) was used to analyze the genes for survival.

## 3. Results

### 3.1. Screening of differential genes

The progression of HBV infection to chronic hepatitis, cirrhosis, and finally to liver cancer is a long-term chronic process. It is the result of multiple mechanisms involving several molecules. Most of the existing studies have analyzed the expression of differential genes in HCC and HBV infected liver tissues to find the mechanism of HBV infection and carcinogenesis.

Elucidating the molecules that change in the early stage of HBV infection through big data analysis, and these changes continue to the HCC stage; then, these molecules may be expected to become early markers for the onset of HCC during the HBV infection stage. These results might provide early clues for possible poor prognosis.

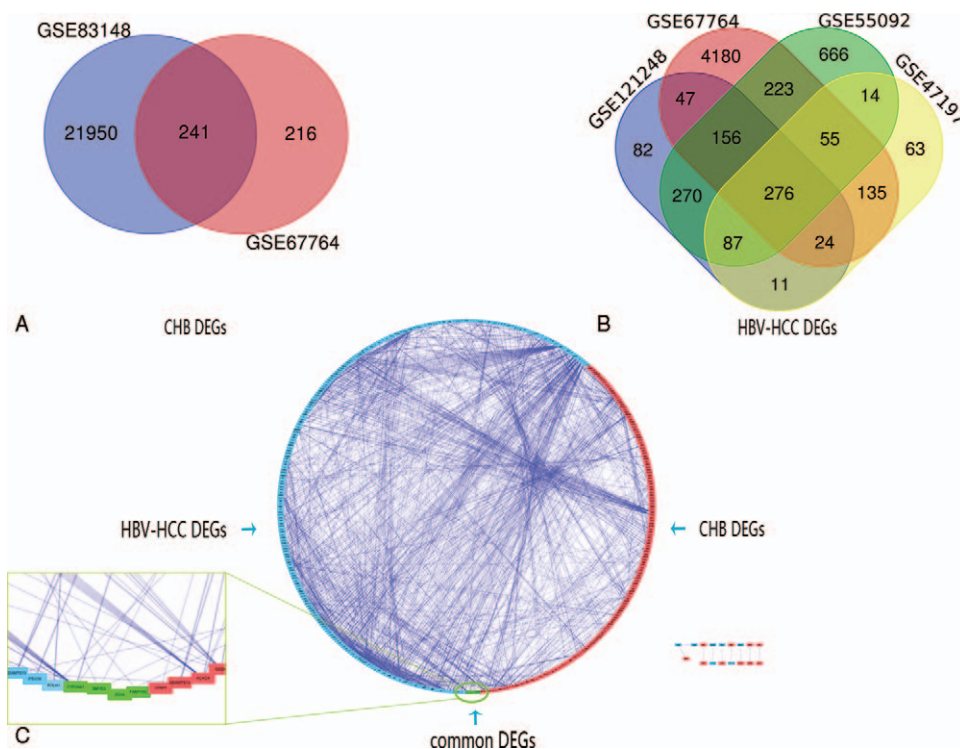
Therefore, we designed 2 sets of transcriptome analysis. The first group analyzed the differences between healthy people and CHB patients (CHB DEGs), and the second group analyzed the differences between CHB patients and HBV-HCC (HBV-HCC DEGs). Fortunately, we obtained the “continuous change” genes that are the “4 common DEGs” by comparing the difference genes between these 2 groups.

In the study, we downloaded GSE83148 and GSE67764 from the GEO database, and used GEO2R to read, preprocess, and screen DEGs. We use fold change to screen DEGs, and the screening criteria are  $P < .05$ ,  $|\log_2FC| \geq 1$ . Consequently, a total of 241 DEGs that were significantly changed in the occurrence of CHB were screened (Fig. 1A), named CHB DEGs. We downloaded the HBV-HCC datasets from the GEO database, including GSE121248, GSE67764, GSE55092, and GSE67764, and used these 4 datasets a total of 276 DEGs that were significantly changed after the occurrence of HBV-HCC were screened (Fig. 1B), named HBV-HCC DEGs.

This study used STRING to analyze CHB DEGs and HBV-HCC DEGs, and obtained PPI interaction diagrams. In addition, 4 genes were found to exist in both CHB DEGs and HBV-HCC DEGs, named 4 common DEGs (Fig. 1C).

### 3.2. The position of the 4 common DEGs in the PPI network

In the above result, we found 4 common DEGs: CYP26A1, ZG16, SMYD3, and FAM110C. The 4 common DEGs changed at the beginning of HBV infection, and lasted during the stage of



**Figure 1.** A: Venn diagram of differential genes in the two datasets of CHB; B: Venn diagram of differential genes in the 4 datasets of HBV-HCC; C: PPI network diagram of CHB DEGs and HBV-HCC DEGs and the discovery of 4 common genes. The blue band represents the HBV-HCC DEGs, the pink represents the CHB DEGs, and the green represents the 4 common DEGs. CHB DEGs=chronic hepatitis B differentially expressed genes, HBV-HCC=hepatitis B associated hepatocellular carcinoma, PPI=protein–protein interaction network.

HBV-HCC. These genes might indicate the occurrence of HBV-HCC from the beginning of HBV infection. Next, we highlighted the position of CYP26A1, ZG16, SMYD3, FAM110C and their closely related gene models (Fig. 2). In this network diagram, we found that CYP26A1, ZG16, SMYD3, and FAM110C were associated with a total of 33 genes, thereby affecting other differential genes in CHB DEGs and HBV-HCC DEGs (Fig. 2). In a word, the 4 common DEGs may play an important role in developing HBV to HCC.

### 3.3. GO and KEGG analysis

Next, to further investigate the functions and mechanisms between the CHB DEGs, HBV-HCC DEGs, CYP26A1, ZG16, SMYD3, FAM110C, and the 33 genes in their models, GO and KEGG enrichment analyses were performed in DAVID.

First, the GO analysis results showed that the CHB DEGs were significantly enriched in cell–cell signaling, signal transduction, ion transmembrane transport, regulation of synaptic vesicle transport, tangential migration from the subventricular zone to the olfactory bulb in the biological processes category. In the cell component analysis, we found that CHB DEGs were predominantly involved in the plasma membrane, extracellular space, proteinaceous extracellular matrix, microvillus. Moreover, in the molecular function analysis, these genes were mainly associated with chemokine activity, GABA-A receptor activity, nucleoside diphosphate kinase activity, and transporter activity. Furthermore, the KEGG pathway enrichment analysis indicated that the CHB DEGs were protein digestion and absorption (hsa04974), neuroactive ligand-receptor interaction (hsa04080), cytokine–cytokine receptor interaction (hsa04060), and retinol metabolism (hsa00830) (Fig. 3A).

Then, the GO analysis results showed that the HBV-HCC DEGs were significantly enriched for oxidation–reduction, complement activation, activation of plasma proteins involved in acute inflammatory response, tryptophan catabolic process, ethanol oxidation in the biological processes category. The cell component analysis revealed that HBV-HCC DEGs were predominantly involved in extracellular space, extracellular region, extracellular region part, and microsome. In the molecular function analysis, these genes were mainly associated with electron carrier activity, oxygen binding, heme binding, and tetrapyrrole binding. Moreover, the KEGG pathway enrichment analysis indicated that the HBV-HCC DEGs were retinol metabolism (hsa00830), complement and coagulation cascades (hsa04610), drug metabolism (hsa00982), tryptophan metabolism (hsa00380), and metabolism of xenobiotics by cytochrome P450 (hsa00980) (Fig. 3B).

Finally, to analyze the processes involved in CYP26A1, ZG16, SMYD3, FAM110C, and the 33 genes in their models for CHB and HBV-HCC, we also performed GO and KEGG analysis. Comparing it with the above CHB DEGs' GO and KEGG results, the results showed that CYP26A1, ZG16, SMYD3, and FAM110C mainly acted on the regulation of synaptic vesicle transport in CHB, tangential migration from the subventricular zone to the olfactory bulb, cellular response to manganese ion, protein localization to mitochondrion, cellular response to dopamine, and negative regulation of neuron death in the biological processes category for CHB DEGs. Moreover, it also acted on retinol metabolism in the KEGG for CHB DEGs (Fig. 3C).

Compared with the above HBV-HCC DEGs' GO and KEGG results, the results showed that CYP26A1, ZG16, SMYD3, and

FAM110C mainly acted on tryptophan catabolic process, ethanol oxidation, drug metabolic process, tryptophan catabolic process to kynurenine, xenobiotic metabolic process, retinoic acid metabolic process, steroid metabolic process, retinoid metabolic process, steroid catabolic process, retinal metabolic process, androgen metabolic process in the biological processes category for HBV-HCC DEGs. Furthermore, it also acted on retinol metabolism, tryptophan metabolism, metabolism of xenobiotics by cytochrome P450, tyrosine metabolism, glycolysis/Gluconeogenesis and linoleic acid metabolism in the KEGG for HBV-HCC DEGs (Fig. 3C).

Based on the above analysis of biological information, we then tried to compare the biological processes and KEGG pathways enriched by CYP26A1, ZG16, SMYD3, and FAM110C in CHB and HBV-HCC. We found that in both CHB and HBV-HCC patients, CYP26A1, ZG16, SMYD3, and FAM110C affect the same KEGG pathway, namely the retinol metabolism pathway (Fig. 4C). The change in the retinol metabolism pathway is the possible reason why CYP26A1, ZG16, SMYD3, and FAM110C change in the early stage of HBV infection, and these changes continue to the HCC stage. Therefore, in addition to the verification of the 4 genes, retinol metabolism in the liver may also be worthy for future exploration of HBV development into HCC.

### 3.4. Survival analysis of CYP26A1, FAM110C, SMYD3, ZG16

Through analyzing the prognosis of these 4 genes in HCC, we found that CYP26A1, FAM110C, SMYD3, and ZG16 are all related to the development of liver cancer and patient survival (Fig. 4A–D). The above research results show that these 4 genes are significantly related to the prognosis of liver cancer patients.

## 4. Discussion

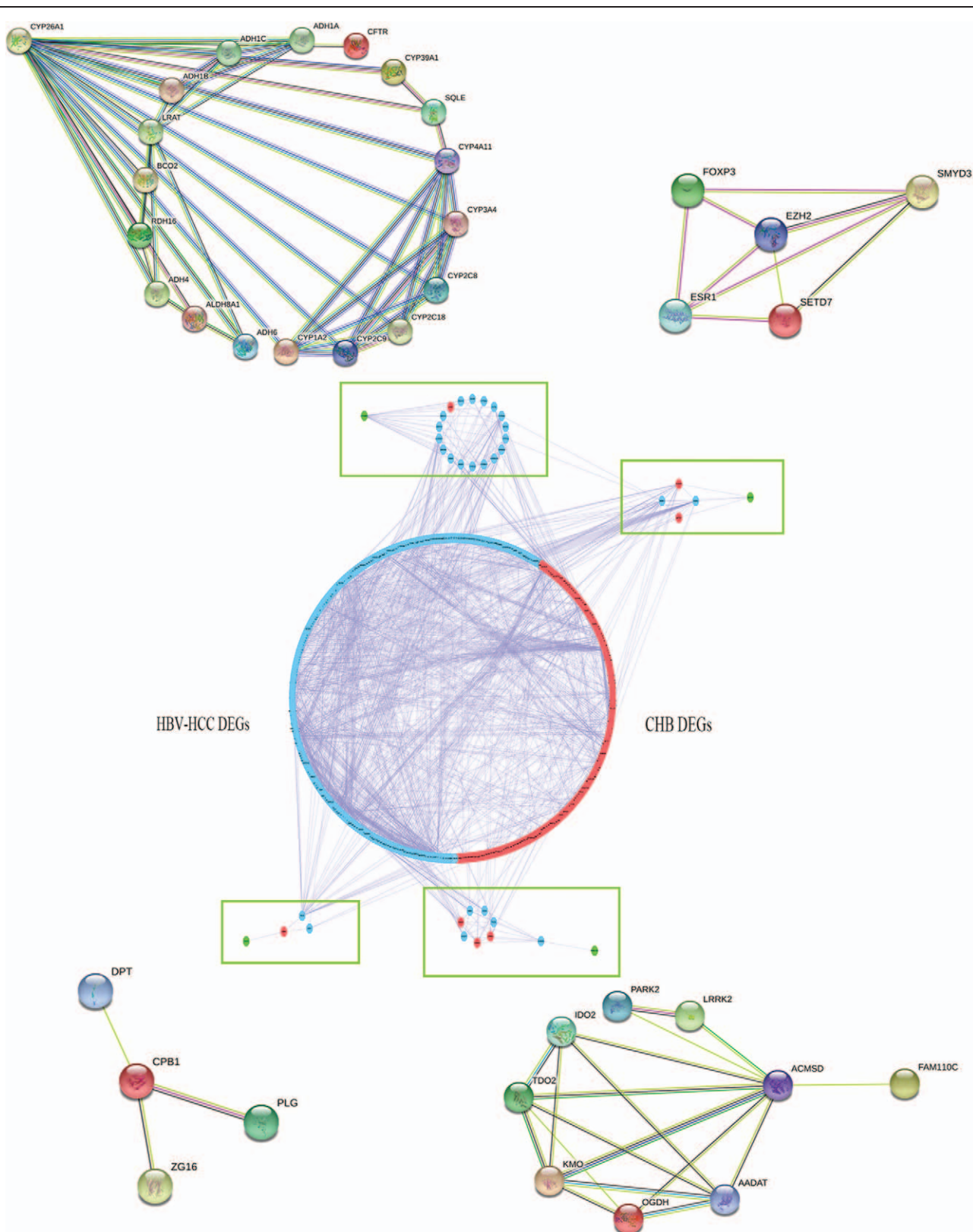
HCC is a serious global health burden. More than half of HCCs in the world are attributed to chronic hepatitis B.<sup>[11]</sup> The relationship between HBV and HCC formation is complicated.<sup>[12]</sup> HBV causes HCC including HBV DNA integration into the host genome,<sup>[13,14]</sup> transactivation of oncogenes by HBV coding proteins, HBV induced chronic inflammation and hepatocyte regeneration.<sup>[15,16]</sup>

The underlying mechanisms of HCC initiation and progression are complicated, numerous studies have been conducted to investigate HCC; however, the timely treatment and early diagnosis of HCC still remain challenging.<sup>[17]</sup> With the multi-omics techniques application, many key molecules related to the occurrence and development of liver cancer have been selected and become targets for early screening and treatment of liver cancer.

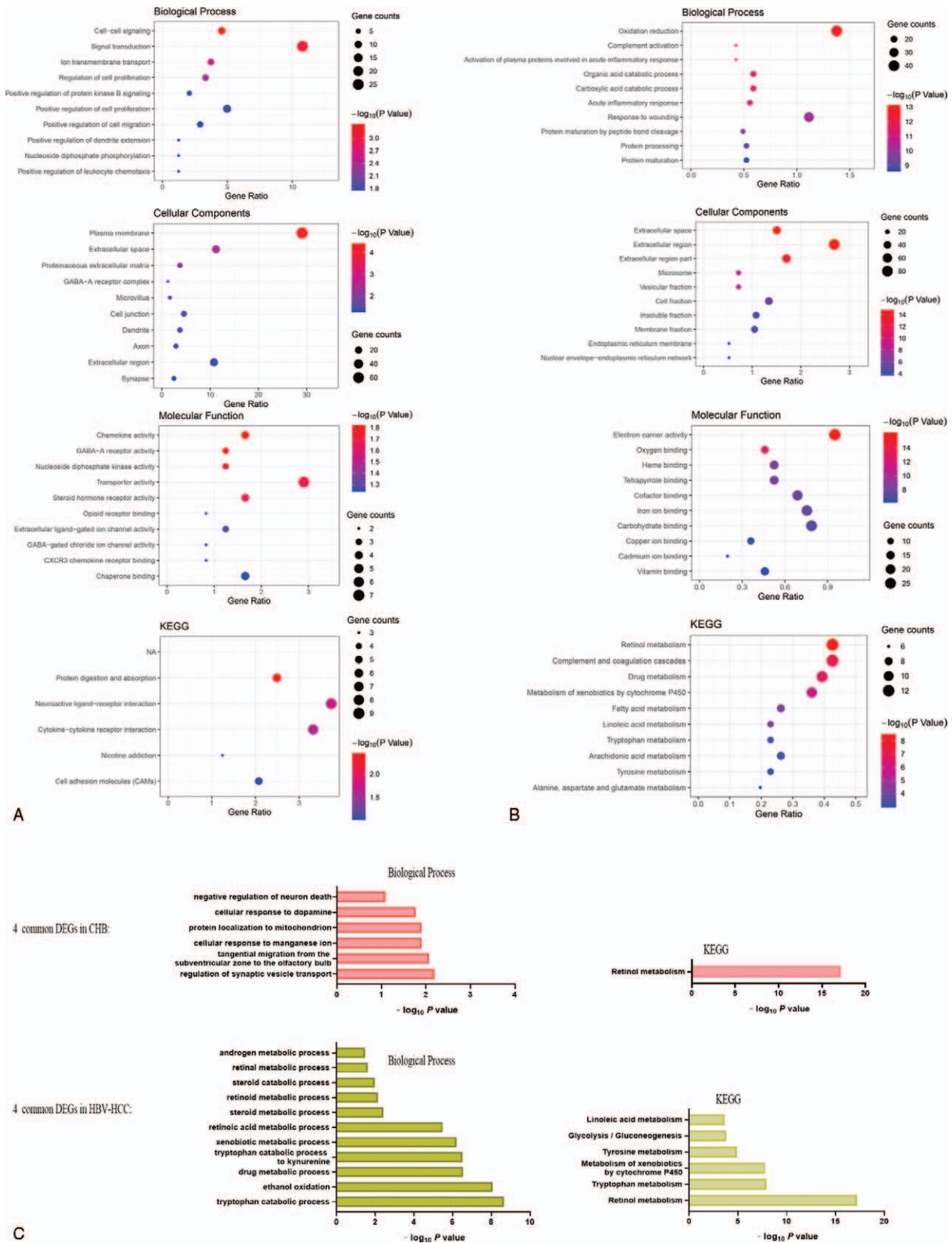
In this study, we screened 241 CHB DEGs in GSE83148, GSE67764, and 276 HBV-HCC DEGs in GSE121248, GSE67764, GSE55092, and GSE55092. CHB DEGs represent the changes in various genes in the patient after HBV infection, and HBV-HCC DEGs represent the changes in various genes in the patient's body after the HBV patient develops further cancer.

By taking the intersection of these two DEGs groups, we found 4 common DEGs: CYP26A1, ZG16, SMYD3, and FAM110C, which changed continuously after HBV infection to HBV-HCC. This result suggested that these 4 genes changed in the early stage





**Figure 2.** The position of CYP26A1, ZG16, SMYD3, FAM110C, and their related gene model in the PPI network diagram of CHB DEGs and HBV-HCC DEGs. The blue band represents the HBV-HCC DEGs, the pink represents the CHB DEGs, and the green represents the 4 common DEGs. CHB DEGs = chronic hepatitis B differentially expressed genes, HBV-HCC = hepatitis B associated hepatocellular carcinoma, PPI = protein-protein interaction network.



**Figure 3.** A: GO enrichment of CHB DEGs; B: GO enrichment of HBV-HCC DEGs; C: The role of 4 common genes in the enrichment results of CHB and HBV-HCC: Enrich the 4 common genes and their PPI models, and the result is compared with the enrichment results of CHB and HBV-HCC. Pink shows the biological processes and KEGG pathway that 4 common genes participate in CHB; green shows the biological processes and KEGG pathway that 4 common genes participate in HBV-HCC. CHB DEGs=chronic hepatitis B differentially expressed genes, HBV-HCC=hepatitis B associated hepatocellular carcinoma, PPI= protein-protein interaction network.

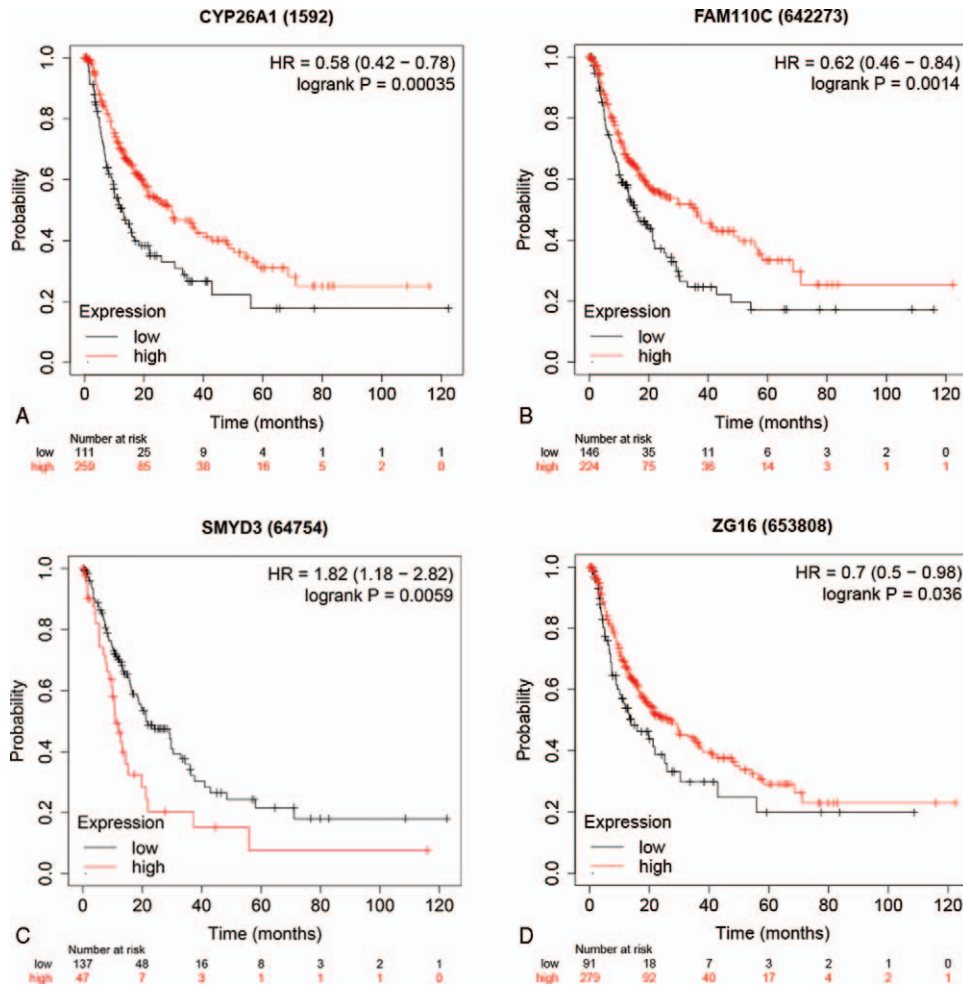


Figure 4. Analysis of the 4 proteins and the survival analysis of liver cancer patients.

of HBV infection and heralded the malignant outcome of the disease.

By GO and KEGG analysis, CHB DEGs were found to play important roles in cell–cell signaling, signal transduction, retinol metabolism, protein digestion, and absorption. HBV-HCC DEGs mainly played important roles in oxidation–reduction, complement activation, retinol metabolism, tryptophan metabolism, and other processes. CYP26A1, ZG16, SMYD3, and FAM110C mainly acted on the regulation of synaptic vesicle transport, negative regulation of neuron death, retinol metabolism in CHB, while CYP26A1, ZG16, SMYD3, and FAM110C mainly acted on tryptophan catabolic process, retinoid metabolic process, and retinol metabolism in HBV-HCC.

According to our further analysis, we found that these 4 genes were closely related to the HCC prognosis.

Regarding these 4 genes, we found that they are more or less related to virus infection, cell proliferation, and tumorigenesis. CYP26A1 is the main scavenger of atRA (all-trans retinoic acid) in the adult liver.<sup>[18]</sup> It is involved in developing various cancers, including breast cancer, cervical cancer, and malignant oral diseases.<sup>[19–21]</sup> It can also participate in EBV virus replication.<sup>[22]</sup> FAM110C participates in cell proliferation and migration, regulates cell proliferation, and regulates reproduction.<sup>[23–25]</sup> SMYD3 has been studied extensively, involving the occurrence

and development of various cancers, including ovarian cancer, breast cancer, rectal cancer, pancreatic cancer, hepatocellular carcinoma, etc.<sup>[26–30]</sup> It is a very valuable anti-cancer target.<sup>[31–33]</sup> In addition, SMYD3 is also involved in the disease process of chronic lymphocytic leukemia and also supports the transcription process of Ebola virus mRNA.<sup>[34]</sup> ZG16 is also involved in the occurrence of various cancers, including colorectal cancer and liver cancer.<sup>[35–37]</sup> Further experiments should be carried out to study the mechanisms of these genes for HBV infection and HCC formation.

Indeed, our research has some potential bias and imprecision. First of all, this kind of reasoning method has never been used before. Statistical methods require deeper thinking and in-depth analysis. The analysis results in this article need to be verified by a large number of basic experiments such as Western Blot and Polymerase Chain Reaction. However, this article aims to provide a new way of thinking to analyze continuous disease specimens. Although this way of thinking is still immature in analysis methods, it does provide new research directions and new hypotheses for basic scientific research and bioinformatics analysis.

In our study, CYP26A1, ZG16, SMYD3, and FAM110C were found to change after HBV infection and played an important role in the occurrence, development, and prognosis of HBV-



HCC. More importantly, these gene changes occur in the early stage of HBV infection, making these genes expected to be an early warning indicator of HBV infection and its poor prognosis. These results also provide a new research direction for the mechanism of HBV-HCC.

### Author contributions

**Conceptualization:** ShanShan Wang, DeXi Chen, Ying Shi.

**Data curation:** Yang Wang, Yang Che.

**Formal analysis:** Yang Wang.

**Investigation:** ShanShan Wang.

**Methodology:** Yang Wang, ShanShan Wang, Yang Che, Ying Shi.

**Resources:** Yang Wang, Yang Che.

**Software:** Yang Wang.

**Supervision:** ShanShan Wang.

**Validation:** Yang Wang.

**Visualization:** Yang Wang.

**Writing – original draft:** Yang Wang, ShanShan Wang, Ying Shi.

**Writing – review & editing:** DeXi Chen, YaLi Liu, Ying Shi.

### References

- Vittal A, Ghany MG. WHO guidelines for prevention, care and treatment of individuals infected with HBV: a US perspective. *Clin Liver Dis* 2019;23:417–32.
- European Association For The Study Of The Liver/EASL clinical practice guidelines: management of chronic hepatitis B virus infection. *J Hepatol* 2012;57:167–85.
- Wang FS, Fan JG, Zhang Z, Gao B, Wang HY. The global burden of liver disease: the major impact of China. *Hepatology* 2014;60:2099–108.
- Gane EJ. Future anti-HBV strategies. *Liver Int* 2017;37:40–4.
- Huang DW, Sherman BT, Tan Q, et al. The DAVID Gene Functional Classification Tool: a novel biological module-centric algorithm to functionally analyze large gene lists. *Genome Biol* 2007;8:R183.
- Gene Ontology Consortium: going forward. *Nucleic Acids Res* 2015;43:D1049–56.
- Kanehisa M, Furumichi M, Tanabe M, Sato Y, Morishima K. KEGG: new perspectives on genomes, pathways, diseases and drugs. *Nucleic Acids Res* 2017;45:D353–61.
- Smoot ME, Ono K, Ruscheinski J, Wang P-L, Ideker T. Cytoscape 2.8: new features for data integration and network visualization. *Bioinformatics* 2011;27:431–2.
- Bandettini WP, Kellman P, Mancini C, et al. MultiContrast Delayed Enhancement (MCODE) improves detection of subendocardial myocardial infarction by late gadolinium enhancement cardiovascular magnetic resonance: a clinical validation study. *J Cardiovasc Magn Reson* 2012;14:83.
- Liu L, Zhao Q, Wei F, et al. Genetic analysis of HIV type 1 env gene in cerebrospinal fluid and plasma of infected Chinese paid blood donors. *AIDS Res Hum Retroviruses* 2012;28:106–9.
- Kulik L, El-Serag HB. Epidemiology and management of hepatocellular carcinoma. *Gastroenterology* 2019;156:477.e1–91.e1.
- Neuveut C, Wei Y, Buendia MA. Mechanisms of HBV-related hepatocarcinogenesis. *Hepatology* 2010;52:594–604.
- Karayiannis P. Hepatitis B virus: virology, molecular biology, life cycle and intrahepatic spread. *Hepatology* 2017;11:500–8.
- Tu T, Budzinska MA, Shackel NA, Urban S. HBV DNA integration: molecular mechanisms and clinical implications. *Viruses* 2017;9:75.
- Furuta M, Tanaka H, Shiraiishi Y, et al. Characterization of HBV integration patterns and timing in liver cancer and HBV-infected livers. *Oncotarget* 2018;9:25075–88.
- Chang TS, Chen CL, Wu YC, et al. Inflammation promotes expression of stemness-related properties in HBV-related hepatocellular carcinoma. *PLoS One* 2016;11:e0149897.
- Tang A, Hallouch O, Chernyak V, Kamaya A, Sirlin CB. Epidemiology of hepatocellular carcinoma: target population for surveillance and diagnosis. *Abdom Radiol (NY)* 2018;43:13–25.
- Thatcher JE, Zelter A, Isoherranen N. The relative importance of CYP26A1 in hepatic clearance of all-trans retinoic acid. *Biochem Pharmacol* 2010;80:903–12.
- Osanai M, Lee GH. The retinoic acid-metabolizing enzyme CYP26A1 upregulates fascin and promotes the malignant behavior of breast carcinoma cells. *Oncol Rep* 2015;34:850–8.
- Wu SJ, Chen YJ, Shieh TY, et al. Association study between novel CYP26 polymorphisms and the risk of betel quid-related malignant oral disorders. *ScientificWorldJournal* 2015;2015:160185.
- Osanai M, Petkovich M. Expression of the retinoic acid-metabolizing enzyme CYP26A1 limits programmed cell death. *Mol Pharmacol* 2005;67:1808–17.
- Jones RJ, Dickerson S, Bhende PM, Delecluse HJ, Kenney SC. Epstein-Barr virus lytic infection induces retinoic acid-responsive genes through induction of a retinol-metabolizing enzyme, DHRS9. *J Biol Chem* 2007;282:8317–24.
- Lou W, Liu J, Ding B, et al. Identification of potential miRNA-mRNA regulatory network contributing to pathogenesis of HBV-related HCC. *J Transl Med* 2019;17:7.
- Hauge H, Fjelland KE, Sioud M, Aasheim HC. Evidence for the involvement of FAM110C protein in cell spreading and migration. *Cell Signal* 2009;21:1866–73.
- Li F, Jang H, Puttabyatappa M, Jo M, Curry TE Jr. Ovarian FAM110C (family with sequence similarity 110C): induction during the periovulatory period and regulation of granulosa cell cycle kinetics in rats. *Biol Reprod* 2012;86:185.
- Giakountis A, Moulos P, Sarris ME, Hatzis P, Talianidis I. SMYD3-associated regulatory pathways in cancer. *Semin Cancer Biol* 2017;42:70–80.
- Lyu T, Jiang Y, Jia N, et al. SMYD3 promotes implant metastasis of ovarian cancer via H3K4 trimethylation of integrin promoters. *Int J Cancer* 2020;146:1553–67.
- Al-Eitan LN, Rababa'h DM. Correlation between a variable number tandem repeat (VNTR) polymorphism in SMYD3 gene and breast cancer: a genotype-phenotype study. *Gene* 2020;728:144281.
- Zhu CL, Huang Q. Overexpression of the SMYD3 promotes proliferation, migration, and invasion of pancreatic cancer. *Dig Dis Sci* 2020;65:489–99.
- Lin F, Wu D, Fang D, Chen Y, Zhou H, Ou C. STAT3-induced SMYD3 transcription enhances chronic lymphocytic leukemia cell growth in vitro and in vivo. *Inflamm Res* 2019;68:739–49.
- Bottino C, Peserico A, Simone C, Caretti G. SMYD3: an oncogenic driver targeting epigenetic regulation and signaling pathways. *Cancers (Basel)* 2020;12:142.
- Rajajeyabalachandran G, Kumar S, Murugesan T, et al. Therapeutic potential of deregulated lysine methyltransferase SMYD3 as a safe target for novel anticancer agents. *Expert Opin Ther Targets* 2017;21:145–57.
- Fenizia C, Bottino C, Corbetta S, et al. SMYD3 promotes the epithelial-mesenchymal transition in breast cancer. *Nucleic Acids Res* 2019;47:1278–93.
- Chen J, He Z, Yuan Y, et al. Host factor SMYD3 is recruited by Ebola virus nucleoprotein to facilitate viral mRNA transcription. *Emerging Microb Infect* 2019;8:1347–60.
- Meng H, Li W, Boardman LA, Wang L. Loss of ZG16 is associated with molecular and clinicopathological phenotypes of colorectal cancer. *BMC Cancer* 2018;18:433.
- Cunningham KA, Bouagnon AD, Barros AG, et al. Loss of a neural AMP-activated kinase mimics the effects of elevated serotonin on fat, movement, and hormonal secretions. *PLoS Genet* 2014;10:e1004394.
- Zhou YB, Cao JB, Yang HM, et al. hZG16, a novel human secreted protein expressed in liver, was down-regulated in hepatocellular carcinoma. *Biochem Biophys Res Commun* 2007;355:679–86.