ChinaRxiv 中X

AI translation · View original & related papers at
chinarxiv.org/items/chinaxiv-202311.00139

Postprint: Identification of Hub Genes in Hepatocellular Carcinoma Based on Weighted Gene Co-expression Network Analysis and TCGA Clinical Data Analysis

Authors: Chen Chao, Chen Tianxiang, Liu Qianwei, Zhang Zhi, Wang Huanhuan, Wu Pingping, Gao Lei, Yu Zhaoxiang, Yu Zhaoxiang

Date: 2023-11-17T00:00:00+00:00

Abstract

Background: Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related death worldwide, accounting for approximately 90% of all primary liver cancer cases. It has high recurrence and mortality rates, and the molecular mechanisms underlying its pathogenesis remain unclear.

Objective: To explore the potential molecular mechanisms of HCC and identify novel biomarkers.

Methods: Gene expression profile GSE62232 was downloaded from the GEO database, and RNA-seq expression data and clinical information were obtained from the TCGA database. Differentially expressed genes between normal liver tissues and HCC tissues were identified through differential gene expression analysis. Enrichment analysis was performed on the differentially expressed genes. Based on the gene expression data profiles of HCC from TCGA and GSE62232, a co-expression network was constructed using the WGCNA R package to conduct Weighted Gene Co-expression Network Analysis (WGCNA). Modules with clinical significance were selected, and candidate hub genes were screened. The candidate hub genes were further analyzed for significant differential expression between HCC tissues and normal liver tissues, as well as for significant correlation with overall survival and disease-free survival of HCC patients, to ultimately identify the hub genes. Protein expression of the hub genes was validated using the Human Protein Atlas database.

Results: The gene expression data in this study were derived from 50 normal liver tissue samples and 373 HCC tissue samples. Differential gene expression analysis identified 7,230 differentially expressed genes between HCC and normal liver tissues (3,691 upregulated genes and 3,539 downregulated genes in

HCC). Enrichment analysis revealed that upregulated differentially expressed genes were primarily involved in cell cycle regulation and mitotic processes, while downregulated differentially expressed genes were mainly involved in small molecule metabolism and organic acid metabolism processes. WGCNA identified 19 gene modules associated with clinical characteristics of HCC patients. By analyzing the relationships between modules and clinical features, the cyan module and purple module were selected. The top two genes in the cyan module that were strongly correlated with both overall survival and disease-free survival of patients were VPS45 and FAM189B. The top two genes in the purple module that were strongly correlated with both overall survival and disease-free survival were CLEC1B and FCN3. Therefore, VPS45, FAM189B, CLEC1B, and FCN3 were identified as the final hub genes. Immunohistochemical staining from the Human Protein Atlas database showed that VPS45 and FAM189B were expressed at higher levels in HCC tissues than in normal liver tissues, FCN3 was expressed at lower levels in HCC tissues than in normal liver tissues, and CLEC1B showed no significant difference in expression between HCC tissues and normal liver tissues.

Conclusion: VPS45, FAM189B, CLEC1B, and FCN3 were preliminarily identified as novel potential biomarkers for HCC. These hub genes may provide a theoretical basis for targeted therapy of HCC.

Full Text

Preamble

Analysis and Identification of Hub Genes in Hepatocellular Carcinoma Based on Weighted Gene Co-expression Network and Cancer Genome Atlas Clinical Data

CHEN Chao[1], CHEN Tianxiang[2], LIU Qianwei[1], ZHANG Zhi[1], WANG Huanhuan[2], WU Pingping[1], GAO Lei[1], YU Zhaoxiang[1]*

[1]Department of General Surgery, The First Affiliated Hospital of Xi' an Medical University, Xi' an 710077, China
[2]Department of Hepatobiliary Surgery, The First Affiliated Hospital of Xi' an Jiaotong University, Xi' an 710061, China

Corresponding author: YU Zhaoxiang, Professor/Chief Physician; E-mail: yuzhaoxiang2@163.com

Abstract
Background: Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related mortality globally, accounting for approximately 90% of all primary liver cancer cases. Its recurrence and mortality rates are high, with the underlying molecular mechanisms remaining unclear.
Objective: To explore potential molecular mechanisms of HCC and identify novel biomarkers.

Methods: Gene expression profile GSE62232 was retrieved from the GEO database, and RNA-seq expression data and clinical information were downloaded from the TCGA database. Differentially expressed genes between normal liver tissue and HCC tissue were identified through differential gene expression analysis, followed by enrichment analysis of these genes. Based on the gene expression data profiles of HCC from TCGA and GSE62232, a co-expression network was established using the WGCNA R package to perform weighted gene co-expression network analysis (WGCNA). Clinically significant modules were selected, and candidate Hub genes were screened. Candidate Hub genes were further analyzed for significant differential expression between HCC and normal liver tissues and for significant correlation with overall survival and disease-free survival in HCC patients, leading to final identification of Hub genes. Protein expression of Hub genes was validated through the Human Protein Atlas database.
Results: The gene expression data in this study were obtained from 50 normal liver tissue samples and 373 HCC tissue samples. Differential gene expression analysis identified 7,230 differentially expressed genes between HCC and normal liver tissue, comprising 3,691 up-regulated genes and 3,539 down-regulated genes in HCC. Enrichment analysis showed that up-regulated differentially expressed genes were mainly involved in cell cycle regulation and mitotic processes, while down-regulated genes participated primarily in small molecule metabolism and organic acid metabolism. WGCNA identified 19 gene modules related to clinical features of HCC patients. By analyzing relationships between modules and clinical features, the cyan and purple modules were selected. In the cyan module, the top two genes strongly associated with both overall survival and disease-free survival were VPS45 and FAM189B. In the purple module, the top two genes strongly associated with both overall survival and disease-free survival were CLEC1B and FCN3. Therefore, VPS45, FAM189B, CLEC1B, and FCN3 were identified as the final Hub genes. Immunohistochemical staining from the Human Protein Atlas database showed that VPS45 and FAM189B were expressed at higher levels in HCC tissues than in normal liver tissues, FCN3 was expressed at lower levels in HCC tissues, and CLEC1B showed no obvious expression difference between HCC and normal liver tissues.
Conclusion: VPS45, FAM189B, CLEC1B, and FCN3 have been preliminarily identified as potential novel biomarkers for HCC. These Hub genes may provide a theoretical foundation for targeted therapy of HCC.

[**Key words**] Hepatocellular carcinoma; Weighted gene co-expression network analysis; Hub gene; Molecular targeted therapy; VPS45; FAM189B; CLEC1B; FCN3

Hepatocellular carcinoma (HCC) is the third leading cause of cancer-related death worldwide and represents the primary cause of cancer-related mortality in China in recent years. HCC accounts for 90% of all primary liver cancer cases. Major risk factors for HCC pathogenesis include chronic hepatitis B virus (HBV)

and hepatitis C virus (HCV) infection, smoking, alcohol consumption, overweight, non-alcoholic fatty liver disease (NAFLD), diabetes, and aflatoxin B1 intake, which cause DNA damage, epigenetic alterations, and cancer-associated mutations that ultimately drive HCC progression. The majority of patients are diagnosed at advanced stages. Although HCC treatment has advanced significantly in recent years, primarily including transcatheter arterial chemoembolization, molecular targeted therapy, ablation, surgical resection, and liver transplantation, the 5-year survival rate remains low, and the 5-year recurrence rate reaches 80-90%, resulting in poor prognosis. The precise mechanisms of HCC development remain unclear, making it urgent and necessary to elucidate the molecular mechanisms of HCC occurrence and progression and to develop new diagnostic and therapeutic approaches to improve clinical outcomes.

Recent advances in high-throughput sequencing technologies have provided novel research methods for characterizing cancer genomics, transcriptomics, and epigenomics. Bioinformatics analysis of expression profiles has been widely used to identify new and more effective potential biomarkers for cancer treatment and patient prognosis. Through big data bioinformatics analysis, multiple large-scale datasets for complex diseases, particularly cancer, can be effectively integrated. Weighted gene co-expression network analysis (WGCNA) is an advanced method with accurate and efficient broad gene analysis advantages, used to construct co-expression modules based on similar gene expression patterns and analyze correlations between clinical trait modules and different genomes. WGCNA has been widely applied to identify clinical trait-related modules and Hub genes in various cancers. For example, one WGCNA-based study linked six Hub genes to progression and prognosis in human renal cell carcinoma, while another identified four Hub genes (TOP2A, CHEK1, TTK, and CENPA) that were highly expressed in aggressive adrenocortical malignancies and significantly negatively correlated with overall survival.

This study combined differential gene expression analysis and WGCNA to analyze mRNA data from HCC in TCGA and GEO databases at both expression and functional levels. Functional enrichment analysis was performed to identify modules associated with HCC patient clinical traits and understand the potential biological functions of these co-expressed genes. Hub genes were identified through these bioinformatics analyses and further validated through survival analysis, expression difference analysis between HCC and normal liver tissues, immunohistochemical staining analysis, and literature review. These results will facilitate clinical understanding of HCC etiology and potential molecular mechanisms and provide new therapeutic targets or biomarkers for HCC.

1. Materials and Methods

1.1 Data Sources and Data Processing

The GSE62232 gene expression profile (corresponding to 81 HCC patients; tumor samples were frozen after hepatectomy at -80°C and compared

with normal liver samples using Affymetrix U133plus v2 arrays [GPL570] for comparative gene expression analysis) was retrieved from the GEO database (http://www.ncbi.nlm.nih.gov/geo/), based on the GPL570 platform ([HG-U133_{Plus}2] Affymetrix Human Genome U133 Plus 2.0 Array). RNA sequencing data (50 normal samples and 373 HCC samples), clinical information, and survival data (59 normal samples and 379 HCC samples) for HCC patients were downloaded from the TCGA database [13] (https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga). Patient clinical information data and sample sizes are shown in Table 1 . Probes were converted to gene symbols according to annotation documentation. Duplicate probes for the same gene were removed by measuring the median expression value of all corresponding probes. Based on these processing results, a total of 11,627 genes were selected for subsequent analysis.

1.2 Screening for Differentially Expressed Genes in HCC

For dimensionality reduction using UMAP, the R package UMAP (version 0.2.7.0) was first used for analysis. The expression profile underwent z-score normalization, followed by UMAP function analysis to obtain the dimensionality-reduced matrix. Limma (linear models for microarray data) [14] is a differential expression screening method based on generalized linear models, used to identify differentially expressed genes between normal liver tissue and HCC tissue. Here, the R package limma (version 3.40.6) was used for differential analysis to obtain differentially expressed genes between normal liver and HCC tissues from the expression profile dataset. The lmFit function was used for multivariate linear regression analysis, followed by eBayes function analysis to ultimately obtain the significance of differential expression for each gene. Differential expression analysis was performed comparing HCC tissue and normal liver tissue to obtain expression datasets. With a fold change of 1.5 and $P<0.05$ considered as significant differences, the lmFit function was further used for multivariate linear regression analysis, followed by eBayes function analysis to ultimately obtain the significance of differential expression for each gene and generate a heatmap of differentially expressed genes.

1.3 Functional Enrichment of Differentially Expressed Genes

For functional enrichment analysis of gene sets, gene GO annotations from org.Hs.eg.db (version 3.1.0) were used as background to map genes to the background set. The R package clusterProfiler (version 3.14.3) was used for enrichment analysis to obtain gene set enrichment results. The minimum gene set was set to 5, maximum to 5,000, with $P<0.05$ and $FDR<0.1$ as significance thresholds.

1.4 Weighted Gene Co-expression Network Analysis

Weighted gene co-expression network analysis is a method for analyzing gene expression patterns across multiple samples, clustering genes with similar expres-

sion patterns and analyzing relationships between modules and specific traits or phenotypes. Based on the gene expression data profiles of HCC from TCGA and GSE62232, the WGCNA R package was used to establish a co-expression network. Using the gene expression profile as an example: first, genes and samples with outlier status were removed using standard deviation methods. The WGCNA method was then used to construct a scale-free co-expression network. Specifically, the first step involved constructing Pearson correlation matrices and average linkage methods for all pairwise genes, followed by using a power function amn=|Cmn|^β (where Cmn is the Pearson's correlation between Gene_m and Gene_n) to construct a weighted adjacency matrix. β is a soft threshold parameter that emphasizes strong correlations between genes. After selecting a power of 5, the adjacency was transformed into a Topological Overlap Matrix (TOM) TOMi,j = (lij + aij)/(min(ki+kj) + 1 - aij). This matrix measures a gene' s network connectivity, defined as the sum of its adjacency with all other genes, for network gene quantification, and calculates corresponding dissimilarities. To classify genes with similar expression profiles into gene modules, average linkage hierarchical clustering was performed based on TOM dissimilarity measures, with a minimum gene tree size of 50. Modules with distance <0.25 were merged, ultimately obtaining 19 co-expression modules, with the grey module considered a collection of genes that could not be assigned to any module.

1.5 Selection of Clinically Significant Modules and Hub Genes in HCC

First, principal component analysis was used to describe module eigengenes, corresponding to a single characteristic expression profile for all genes within each module. Correlations between clinical traits and these eigengenes were calculated to identify modules with clinical significance. The linear relationship between gene expression and clinical traits was assigned a gene significance equal to the log of the individual gene P-value. If gene significance was closely related to module membership, defined as the correlation between the module' s eigengene and individual gene expression profiles, it was concluded that the module' s central genes were associated with HCC [15], and these genes were considered candidate Hub genes. As shown in previous studies [16], gene expression correlations were calculated to obtain gene significance, while correlations between module eigengenes and gene expression were calculated to obtain module membership. According to cutoff criteria, 76 genes with high connectivity in clinically significant modules were identified as candidate Hub genes.

1.6 Gene Validation and Bioinformatics Validation of Hub Genes

The GEPIA online website (http://gepia.cancer-pku.cn/) was used to analyze Hub gene expression levels in HCC samples, and Kaplan-Meier analysis was performed using the survival package in R (version 3.2-7). First, differential gene expression profiles and prognosis data from 359 HCC tumor samples were obtained from TCGA. The median expression value for each gene was deter-

mined, and samples were assigned to "high expression" or "low expression" groups for a given gene based on whether expression was above or below the median. The log-rank test was used to evaluate significance in overall survival and disease-free survival between high and low expression groups. Genes with $P<0.05$ were considered validated Hub genes. Then, based on data from TCGA and the GEPIA website, differences in Hub gene expression between normal liver tissue and HCC tissue were screened. Expression levels were normalized by mean, and differences associated with $P<0.05$ were considered statistically significant.

1.7 Validation of Hub Gene Protein Expression Using the Human Protein Atlas Database

The Human Protein Atlas database (https://www.proteinatlas.org/) primarily provides tissue and cellular distribution information for various human proteins. Immunohistochemical staining results from this database were used to validate protein expression of survival-related genes between HCC tissue and normal liver tissue.

2. Results

2.1 Data Preprocessing

Expression data in this study were obtained from 50 normal liver tissue samples and 373 HCC tissue samples. Based on principal component analysis, 14 tumor samples were excluded from the data (Figure 1 [Figure 1: see original paper]A). Gene expression profiles from the remaining 409 samples were used for subsequent analysis.

2.2 Identification of Differentially Expressed Genes in HCC Samples and GO Enrichment Analysis

A total of 7,230 differentially expressed genes were identified between 50 normal samples and 359 HCC samples. The volcano plot of differentially expressed genes is shown in Figure 1B, including 3,691 up-regulated genes and 3,539 down-regulated genes. Differential analysis results are summarized in Table 2. To explore the potential biological functions of differentially expressed genes in HCC, GO enrichment analysis was performed. Up-regulated differentially expressed genes were mainly involved in cell cycle regulation, mitotic processes, nuclear division, and chromosome segregation (Figure 2 [Figure 2: see original paper]A, 2C). In contrast, down-regulated differentially expressed genes were mainly involved in response to external stimuli, small molecule metabolism, and organic acid metabolism (Figure 2B, 2D).

2.3 WGCNA and Identification of Key Modules

Weighted gene co-expression network analysis was performed based on the expression matrix of 7,230 differentially expressed genes and clinical data from 409 HCC samples. Clustering analysis was conducted to examine data quality across 409 samples, showing that all samples clustered within cutoff thresholds. Six clinical variables were applied in the weighted gene co-expression network analysis: disease status (Tumor or Normal), sex, age, weight, TNM stage, and microvascular invasion (Figure 3 [Figure 3: see original paper]A). The 409 samples were divided into two clusters: tumor and normal. A cluster dendrogram of clinical features and data for HCC samples is shown (Figure 3B).

To construct a scale-free network, the soft threshold β was set to 5, independence to 0.86, and average connectivity approached 0 (Figure 4 [Figure 4: see original paper]A, 4B). Module parameters were set with a minimum module size of 30, sensitivity of 3, and module merge threshold of 0.25. Differentially expressed genes with similar expression patterns were clustered into the same module. Additionally, modules with distance <0.25 were merged, ultimately obtaining 19 co-expression modules. A heatmap of module eigengenes was generated (Figure 4C), with the grey module representing a collection of genes that could not be assigned to any module.

Relationships between modules and clinical traits were then evaluated. The cyan module eigengene showed strong positive correlation with HCC (cor=0.64, P=3.6$\times10^{-50}$), while the purple module eigengene showed strong negative correlation with HCC (cor $= -0.80$, $P = 5.6\times10^{-97}$). These results suggest the cyan module may promote HCC tumorigenesis, while the purple module may protect against HCC. Therefore, Hub genes in the cyan and purple modules were analyzed.

2.4 Identification of Candidate Hub Genes from Cyan and Purple Modules

The criteria for selecting Hub genes in the cyan module were relatively lower than the standard cutoff threshold (MM>0.8). In the cyan module, the top 8 genes meeting thresholds of cor.gene Module Membership > 0.7 and cor.gene Trait Significance > 0.5 were selected: TOMM40L, VPS45, MSTO1, FAM189B, TTC13, PYGO2, NVL, and EHMT2. In the purple module, the top 16 genes meeting thresholds of cor.gene Module Membership > 0.8 and cor.gene Trait Significance > 0.7 were selected: CLEC4M, BMP10, CLEC1B, CLEC4G, GDF2, NDST3, BMPER, STAB2, CCL23, CHRM2, COL6A6, CRHBP, FCN3, and CCBE1 (Figure 5 [Figure 5: see original paper]).

2.5 Hub Gene Expression and Its Correlation with Survival

Based on expression data and clinical information for tumor samples from the GEPIA2 online database, potential associations between gene expression and patient survival were analyzed for genes identified in the cyan and purple modules (Figure 6 [Figure 6: see original paper]). In the cyan module, the top 3

genes strongly associated only with overall survival were TOMM40L, VPS45, and FAM189B; the top 2 genes strongly associated with both overall survival and disease-free survival were VPS45 and FAM189B. In the purple module, the top 3 genes strongly associated only with overall survival were CLEC1B, CCL23, and FCN3; the top 2 genes strongly associated with both overall survival and disease-free survival were CLEC1B and FCN3. Therefore, VPS45, FAM189B, CLEC1B, and FCN3 were identified as the final Hub genes. Genes showing significant differences in disease-free survival in both modules included VPS45, FAM189B, CLEC1B, FCN3, and BMPER.

Using the GEPIA website, expression of these module genes between HCC tissue and normal tissue was analyzed (Figure 7 [Figure 7: see original paper]). Results showed VPS45 and FAM189B were up-regulated in HCC tissues, while CLEC1B and FCN3 were down-regulated. Similar results were obtained using data from the GEO database.

2.6 Immunohistochemical Staining Validation

Immunohistochemical staining from the Human Protein Atlas (HPA) database [17] (available at: Human Protein Atlas proteinatlas.org) showed that VPS45 and FAM189B were expressed at higher levels in HCC tissues than in normal liver tissues, while FCN3 was expressed at lower levels in HCC tissues. CLEC1B showed no obvious expression difference between HCC and normal liver tissues (Figure 8 [Figure 8: see original paper]).

3. Discussion

HCC screening involves regular examination of patients at high risk for HCC, with the goal of detecting HCC at an early stage and providing timely intervention to improve patient survival and quality of life. In clinical applications for HCC, common traditional screening methods include liver ultrasound and serum alpha-fetoprotein, which have relatively limited sensitivity and specificity. Therefore, developing new and more accurate detection methods is of significant clinical value. Current research has identified several genes as novel biomarkers for HCC diagnosis. For example, long non-coding RNA KDM4A-AS1 is overexpressed in many tumors, especially HCC, with levels positively correlated with disease stage and tumor grade and negatively correlated with patient survival [18]. LV et al. [19] showed that BAI3 and CKAP2L may be potential prognostic factors and therapeutic targets for colorectal cancer.

In this study, differentially expressed genes and weighted gene co-expression network analysis were combined to enhance the ability to identify HCC-related genes. Functional enrichment of differential genes revealed that genes mainly involved in cell cycle regulation were dysregulated in HCC tissues, consistent with recent research findings [20]. Disruption of normal cell cycle is a cause of cancer, and targeting cancer cell cycle regulation is a potential therapeutic approach [21]. Bioinformatics analysis based on TCGA and GEO databases

identified important modules and Hub genes in HCC. Two key modules were identified: the cyan module and the purple module. Correlations between genes in these modules and patient overall survival and disease-free survival were analyzed. The top two significantly correlated genes from the cyan module, VPS45 and FAM189B, and the top two from the purple module, CLEC1B and FCN3, were selected as the final Hub genes.

VPS45-deficient neutrophils and fibroblasts show reduced $ $1 integrin levels on their surface, impaired motility, and increased apoptosis in VPS45-deficient fibroblasts [22]. Additionally, studies have shown VPS45 expression levels are associated with various cancers. For example, YAMANOI et al. [23] demonstrated that VPS45 expression levels are associated with malignant ovarian cancer. FAM189B, also known as COTE1, can physically associate with tumor suppressor domain oxidoreductase through mechanism analysis, and is closely related to HCC cell invasion [24]. Overexpression of FAM189B protein and mRNA may increase gastric cancer incidence, with the cell cycle pathway being the most significantly enriched pathway in its KEGG analysis [25]. CLEC1B is a C-type lectin domain family 1 member B associated with immune infiltration in hepatocellular carcinoma, and overexpressed CLEC1B inhibits proliferation and migration of HuH7 cells [26]. FCN3 is a secreted lectin that can activate the complement pathway. Ectopic expression of FCN3 can activate endoplasmic reticulum stress unfolded protein response, inhibit endoplasmic reticulum stress, and improve survival of lung adenocarcinoma cells. FCN3 functions as a tumor suppressor by inducing endoplasmic reticulum stress [27].

This study has limitations inherent to data mining. The data used were obtained from online databases, and sample data may have bias, with sequencing results potentially having technical deviations. To improve reliability of WGCNA results, immunohistochemical data from the Human Protein Atlas were used for confirmation. However, due to database limitations, it was not possible to obtain all relevant IHC data for tumors and adjacent normal samples for each gene. The molecular regulatory mechanisms of the identified Hub genes in HCC prognosis require further validation through clinical data and basic experiments. These genes may serve as molecular markers for early HCC screening, prognostic markers, and therapeutic targets for HCC, providing assistance for population screening or HCC patient treatment.

In summary, this study performed comprehensive bioinformatics analysis to identify potential predictive biomarkers between HCC and normal liver tissues. The results suggest that VPS45, FAM189B, CLEC1B, and FCN3 may be novel potential biomarkers for HCC with special clinical significance. However, validation in large clinical samples is needed, and further exploration is warranted in the future.

Author Contributions: CHEN Chao conceived the research objectives, performed data analysis, and wrote the manuscript. CHEN Tianxiang performed data validation. LIU Qianwei, ZHANG Zhi, WANG Huanhuan, GAO Lei, and WU Pingping were responsible for data curation. YU Zhaoxiang was responsi-

ble for quality control and review of the article, had overall responsibility for the article, and provided supervision and management.

Conflict of Interest: The authors declare no conflict of interest.

References

[1] SUNG H, FERLAY J, SIEGEL R L, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries[J]. CA Cancer J Clin, 2021, 71(3): 209-249. DOI: 10.3322/caac.21660.

[2] MEHTA A, HERRERA H, BLOCK T. Glycosylation and liver cancer[J]. Adv Cancer Res, 2015, 126: 257-279. DOI: 10.1016/bs.acr.2014.11.005.

[3] VOGEL A, MEYER T, SAPISOCHIN G, et al. Hepatocellular carcinoma[J]. Lancet, 2022, 400(10360): 1345-1362. DOI: 10.1016/s0140-6736(22)01200-4.

[4] Cancer Genome Atlas Research Network Electronic Address. Comprehensive and integrative genomic characterization of hepatocellular carcinoma[J]. Cell, 2017, 169(7): 1327-1341.e23. DOI: 10.1016/j.cell.2017.05.046.

[5] LLOVET J M, REAL M I, MONTAÑA X, et al. Arterial embolisation or chemoembolisation versus symptomatic treatment in patients with unresectable hepatocellular carcinoma: a randomised controlled trial[J]. Lancet, 2002, 359(9319): 1734-1739. DOI: 10.1016/S0140-6736(02)08649-X.

[6] SIEGEL R L, MILLER K D, JEMAL A. Cancer statistics, 2017[J]. CA Cancer J Clin, 2017, 67(1): 7-30. DOI: 10.3322/caac.21387.

[7] XIA F, WU L L, LAU W Y, et al. Adjuvant sorafenib after heptectomy for Barcelona clinic liver cancer-stage C hepatocellular carcinoma patients[J]. World J Gastroenterol, 2016, 22(23): 5384-5392. DOI: 10.3748/wjg.v22.i23.5384.

[8] LI N, ZHAN X Q. Identification of clinical trait-related lncRNA and mRNA biomarkers with weighted gene co-expression network analysis as useful tool for personalized medicine in ovarian cancer[J]. EPMA J, 2019, 10(3): 273-290. DOI: 10.1007/s13167-019-00175-0.

[9] KUENZI B M, IDEKER T. A census of pathway maps in cancer systems biology[J]. Nat Rev Cancer, 2020, 20(4): 233-246. DOI: 10.1038/s41568-020-0240-7.

[10] LANGFELDER P, HORVATH S. WGCNA: an R package for weighted correlation network analysis[J]. BMC Bioinformatics, 2008, 9: 559. DOI: 10.1186/1471-2105-9-559.

[11] YUAN L S, CHEN L, QIAN K Y, et al. Co-expression network analysis identified six hub genes in association with progression and prognosis in human clear cell renal cell carcinoma (ccRCC)[J]. Genom Data, 2017, 14: 132-140. DOI: 10.1016/j.gdata.2017.10.006.

[12] XIA W X, YU Q, LI G H, et al. Identification of four hub genes associated with adrenocortical carcinoma progression by WGCNA[J]. PeerJ, 2019, 7: e6555. DOI: 10.7717/peerj.6555.

[13] COLAPRICO A, SILVA T C, OLSEN C, et al. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data[J]. Nucleic Acids Res, 2016, 44(8): e71. DOI: 10.1093/nar/gkv1507.

[14] RITCHIE M E, PHIPSON B, WU D, et al. Limma powers differential expression analyses for RNA-sequencing and microarray studies[J]. Nucleic Acids Res, 2015, 43(7): e47. DOI: 10.1093/nar/gkv007.

[15] DI Y, CHEN D S, YU W, et al. Bladder cancer stage-associated hub genes revealed by WGCNA co-expression network analysis[J]. Hereditas, 2019, 156: 7. DOI: 10.1186/s41065-019-0083-y.

[16] TANG J N, KONG D G, CUI Q X, et al. Prognostic genes of cancer identified by gene co-expression network analysis[J]. Front Oncol, 2018, 8: 374. DOI: 10.3389/fonc.2018.00374.

[17] KARLSSON M, ZHANG C, MÉAR L, et al. A single-cell type transcriptomics map of human tissues[J]. Sci Adv, 2021, 7(31): eabh2169. DOI: 10.1126/sciadv.abh2169.

[18] CHEN T X, LIU R K, NIU Y S, et al. HIF-1α-activated long non-coding RNA KDM4A-AS1 promotes hepatocellular carcinoma progression via the miR-411-5p/KPNA2/AKT pathway[J]. Cell Death Dis, 2021, 12(12): 1152. DOI: 10.1038/s41419-021-04456-8.

[19] LV Y M, XIE B B, BAI B J, et al. Weighted gene coexpression analysis indicates that PLAGL2 and POFUT1 are related to the differential features of proximal and distal colorectal cancer[J]. Oncol Rep, 2019, 42(6): 2473-2485. DOI: 10.3892/or.2019.7368.

[20] SCHAFER K. The cell cycle: a review[J]. Vet Pathol, 1998, 35: 461-478. DOI: 10.1177/030098589803500601.

[21] AARTS M, LINARDOPOULOS S, TURNER N C. Tumour selective targeting of cell cycle kinases for cancer treatment[J]. Curr Opin Pharmacol, 2013, 13(4): 529-535. DOI: 10.1016/j.coph.2013.03.012.

[22] VILBOUX T, LEV A, MALICDAN M C V, et al. A congenital neutrophil defect syndrome associated with mutations in VPS45[J]. N Engl J Med, 2013, 369(1): 54-65. DOI: 10.1056/NEJMoa1301296.

[23] YAMANOI K, BABA T, ABIKO K, et al. Acquisition of a side population fraction augments malignant phenotype in ovarian cancer[J]. Sci Rep, 2019, 9(1): 14215. DOI: 10.1038/s41598-019-50794-w.

[24] ZHANG H, HUANG C J, TIAN Y, et al. Ectopic overexpression of COTE1 promotes cellular invasion of hepatocellular carcinoma[J]. Asian Pac J Cancer Prev, 2012, 13(11): 5799-5804. DOI: 10.7314/apjcp.2012.13.11.5799.

[25] WU C L, TAN Q X, LIU D, et al. High FAM189B expression and its prognostic value in patients with gastric cancer[J]. Biomed Res Int, 2021, 2021: 8875971. DOI: 10.1155/2021/8875971.

[26] ZHANG G H, SU L S, LV X P, et al. A novel tumor doubling time-related immune gene signature for prognosis prediction in hepatocellular carcinoma[J]. Cancer Cell Int, 2021, 21(1): 522. DOI: 10.1186/s12935-021-02227-w.

[27] JANG H, JUN Y, KIM S, et al. FCN3 functions as a tumor suppressor of lung adenocarcinoma through induction of endoplasmic reticulum stress[J]. Cell Death Dis, 2021, 12(4): 407. DOI: 10.1038/s41419-021-03675-y.

(Received: March 17, 2023; Revised: August 18, 2023)
(Editor: JIA Mengmeng)

Note: Figure translations are in progress. See original paper for figures.

Source: ChinaXiv —Machine translation. Verify with original.