European Review for Medical and Pharmacological Sciences 2023; 27: 9586-9600

Potential functions and causal associations of *VPS29* in hepatocellular carcinoma: a bioinformatic and Mendelian randomization study

Y.-Q. YAO[1], J.-L. LV[2]

[1]Guangxi University of Chinese Medicine, Nanning, China
[2]The First Affiliated Hospital of Guangxi University of Chinese Medicine, Nanning, China

Abstract. – OBJECTIVE: Vacuolar protein sorting-associated protein 29 (*VPS29*) plays a certain role in cancer, but its biological significance in hepatocellular carcinoma (HCC) has not been studied. We utilized bioinformatics and Mendelian randomization (MR) analysis to explore the potential function of the *VPS29* gene in HCC, as well as the causal relationship between *VPS29* protein and HCC.

MATERIALS AND METHODS: We downloaded the raw data from TCGA, GEO, and IEU OpenGWAS databases. We used R software for data processing and analysis to explore the relationship between the *VPS29* gene and the expression, prognosis, clinical features, methylation, immune microenvironment, tumor mutation burden, and drug sensitivity in HCC patients. Additionally, a two-sample Mendelian randomization analysis was conducted to investigate the causal relationship between the *VPS29* protein and HCC.

RESULTS: *VPS29* was found to be overexpressed in various types of cancer, including HCC, and its elevated expression often predicts poor prognosis in HCC patients. Univariate and multivariate Cox analysis demonstrated that *VPS29* was an independent prognostic factor in HCC patients. The ROC curve indicated that *VPS29* has a high diagnostic value in HCC. There were differential expressions of *VPS29* in various clinical feature subgroups. The expression of *VPS29* was negatively correlated with methylation levels, and multiple methylation sites were identified in the promoter region, including cg20877181, cg03867797, cg10025392, cg21605021, which were associated with poorer overall survival (OS) at low methylation levels. *VPS29* was associated with immune cell infiltration disorders, including CD8+ T cells, Eosinophils, Neutrophils, Tcm, NK CD56bright cells, TFH, Th2 cells, Th17 cells, etc. Drug sensitivity analysis showed that *VPS29* could be indicative of treatment response to 10 common antineoplastic drugs in different expression subgroups. Inverse variance weighted (IVW) analysis revealed a significant increase in HCC risk associated with *VPS29* [odds ratio (OR): 1.440; 95% confidence interval (CI): 1.195-1.736], and sensitivity analysis showed no heterogeneity or pleiotropy.

CONCLUSIONS: *VPS29* is a risk factor for the occurrence and progression of HCC and may serve as a molecular biomarker for the diagnosis and prognosis of HCC.

Key Words:
Vacuolar protein sorting-associated protein 29 (*VPS29*), Two-sample Mendelian randomization study, Genome-wide association study, Hepatocellular carcinoma, Bioinformatics.

Introduction

Hepatocellular carcinoma (HCC), the most common type of liver cancer, is a malignancy with significant heterogeneity and strong aggressiveness[1]. It results from an interplay between multiple factors, such as viruses and ingestion factors[2,3], and commonly presents insidiously at onset[4]. Despite the continuous improvement in treatment strategies for HCC, the 5-year survival rate for HCC patients worldwide remains discouraging[4,5]. Patients suffering from HCC generally have a poor prognosis, which is a critical difficulty in current clinical treatment. To improve the prognosis of patients with HCC, it is necessary to look for effective and reliable signature genes that may contribute to clinical treatment.

Vacuolar protein sorting-associated protein 29 (*VPS29*) belongs to a group of vacuolar protein sorting genes. The protein encoded by *VPS29* is a component of a large multimeric complex. VPS29 protein is a component of both retromer (VPS26/VPS29/VPS35) and retriever (DSCR3/VPS29/C16orf62) complexes. These are two distinct yet interconnected complexes that, in con-

Corresponding Author: Jianlin Lv, MD; e-mail: lvjl@gxtcmu.edu.cn

junction with the CCDC22/CCDC93/COMMD (CCC) complex, regulate the endosome-to-plasma membrane and endosome-to-trans-Golgi network (TGN) recycling of transmembrane cargo[6-8]. VPS29 protein, being an integral part of the retromer and retriever complexes, is likely to have a central role in coordinating the activity and communication between various retrieval machines by interacting with cooperating proteins[6]. Current studies[9-12] suggest significant roles for the retromer complex in growth factor receptor trafficking and intracellular signaling, both factors having potential implications on their oncogenic capabilities. Among them, the *VPS35* gene has been found to be associated with HCC[13], gastric cancer[14], and colorectal cancer[14]. Within hepatocytes, the VPS35 protein regulates the transport of receptor tyrosine kinase fibroblast growth factor receptor 3 (FGFR3), and changes in its mutation and expression levels can affect cell proliferation through the phosphatidylinositol 3-kinase (PI3K)/AKT signaling pathway[13]. This literature suggests that subunits of the retromer complex may disrupt cell signaling and drive cancer progression, and the role of *VPS29* in cancer is yet to be fully explored.

In recent years, genome-wide association studies (GWAS) have accumulated millions of datasets that link genetic variations to complex diseases or phenotypes[15]. Mendelian randomization (MR) analysis utilizes genetic variations as instrumental variables for exposure, which can strengthen causal inferences of exposure-outcome associations by minimizing confounding factors and reverse causality[16]. In this study, a comprehensive bioinformatics approach combined with MR analysis was employed to deeply investigate the expression, biological functions, prognostic role, and diagnostic value of the *VPS29* gene in HCC, as well as to assess the causal relationship between VPS29 protein and HCC. The flowchart of this study is shown in Figure 1.

Materials and Methods

Data Collection and Pan-Cancer Analysis

The expression levels of *VPS29* in 33 cancer types were obtained from the TCGA database (https://tcga-data.nci.nih.gov/tcga/). We acquired mRNA expression data, clinical information, methylation data, and tumor mutation data (including 374 HCC tissue samples and 50 normal liver tissue samples) from TCGA-LIHC. The expression levels of *VPS29* in HCC were further validated using multiple independent

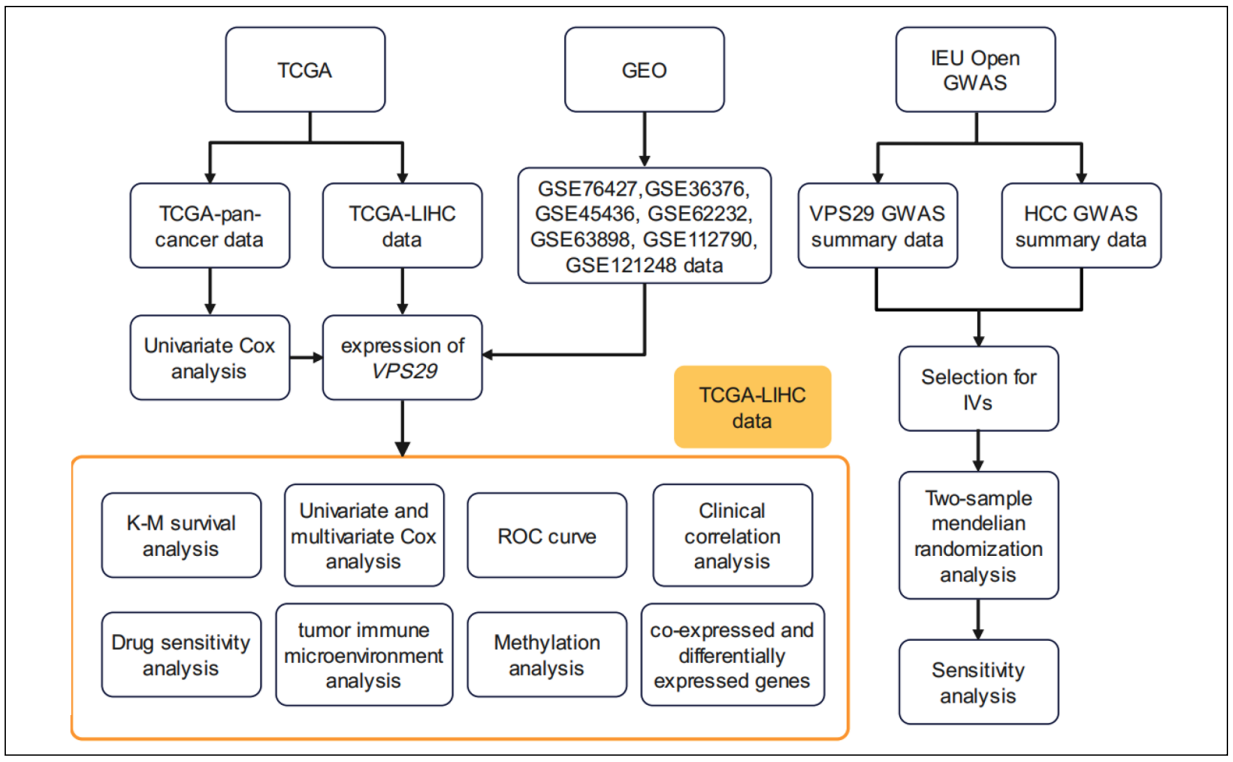

Figure 1. Process flowchart.

datasets downloaded from the GEO database (https://www.ncbi.nlm.nih.gov/geo/), including GSE76427, GSE36376, GSE45436, GSE62232, GSE63898, GSE112790, and GSE121248. The GWAS summary data of VPS29 and HCC were obtained from the integrative epidemiology unit (IEU) Open GWAS database (https://gwas.mrcieu.ac.uk/). The study populations were of European ancestry, and the search codes were "prot-a-3203" and "finn-b-C3_LIVER_INTRAHEPATIC_BILE_DUCTS". The differential expression of *VPS29* in different cancer tissues and their corresponding normal tissues was compared. Univariate Cox analysis was performed to identify the impact of *VPS29* expression levels on the survival of different cancer patients. R packages "ggpubr" and "orestplot" were used to visualize the results through boxplots and forest plots.

The Prognostic and Diagnostic Value of VPS29 in HCC

The differential expression of *VPS29* in HCC and normal tissues was validated through multiple independent datasets. Based on the TCGA-LIHC dataset, HCC patients were divided into a high *VPS29* expression group and a low *VPS29* expression group. Kaplan-Meier (K-M) plots were conducted using the "survival" and "survminer" packages to investigate the relationship between *VPS29* expression and the prognosis [overall survival (OS), disease-specific survival (DSS), and progress-free interval (PFI)] of HCC patients. The "timeROC" package in R was used to plot receiver operating characteristic (ROC) curves to assess the sensitivity and accuracy of *VPS29* in predicting HCC prognosis and diagnosis. Univariate and multivariate Cox regression analyses were performed to evaluate whether *VPS29* is an independent factor in predicting OS in HCC patients. Boxplots were generated to demonstrate the correlation between *VPS29* and different clinical subtypes.

Methylation Analysis

We employed Spearman correlation analysis to evaluate the relationship between the methylation levels of the *VPS29* promoter region and its mRNA expression levels. Liver cancer patients were divided into a high methylation group and a low methylation group according to the median methylation level of *VPS29*. K-M survival analysis was performed to assess the prognostic value of *VPS29* methylation sites.

Drug Sensitivity, Tumor Immune Microenvironment (TIME), and Tumor Mutation Burden (TMB) Analysis

Using the R package "oncoPredict", we predicted the sensitivity of HCC patients to commonly used anticancer drugs. The ridge regression method was employed to calculate the half-maximal inhibitory concentration (IC50) of the drugs, and the difference in IC_{50} between the high and low *VPS29* expression groups was compared. The TIDE database (http://tide.dfci.harvard.edu/) was used to analyze the immune therapy scores of each patient in the TCGA-LIHC dataset, evaluating the response to immunotherapeutic drugs in the high and low *VPS29* expression groups. The infiltration of 24 immune cell types in HCC samples was assessed using the single-sample gene set enrichment analysis (ssGSEA) algorithm provided by the R package "GSVA" (Barcelona, Catalonia, Spain). Since TMB is associated with immune therapy response, we used Spearman correlation analysis to evaluate the relationship between TMB scores and *VPS29* expression.

Screening Co-Expressed Genes and Differentially Expressed Genes

We identified genes that were co-expressed with *VPS29* ($|R| > 0.6$, $p < 0.001$). Univariate regression analysis was conducted to explore the prognostic impact of these co-expressed genes. The R package "limma" (Parkville, Victoria, Australia) was used to identify differentially expressed genes (DEGs) between the high and low *VPS29* expression groups, with $|\log_2 FC| > 1$ and $FDR < 0.05$ considered as significant criteria.

Gene Enrichment Analysis

We used the R package "clusterProfiler" (Guangzhou, China) to conduct gene ontology (GO), Kyoto Encyclopedia of Genes and Genomes (KEGG), and gene set enrichment analysis (GSEA) on DEGs, with the Hallmark v7.2 gene set used for the GSEA in this study.

Two-Sample Mendelian Randomization Analysis

To meet the three basic assumptions for using single nucleotide polymorphisms (SNP) as IV in MR analysis, the following criteria were employed to select IV: (1) SNP associated with VPS29 protein at the genome-wide significance level ($p < 5 \times$E-06). (2) Exclusion of SNP in linkage disequilibrium (LD) (within a 10,000 kb window, $r^2 > 0.001$) to ensure the independent

contribution of the selected SNP. (3) Removal of SNP with an F-statistic less than 10 to ensure sufficiently strong instrumental variables. (4) Exclusion of SNP that are significantly associated with the outcome ($p < 5 \times$ E-08). (5) Harmonization of all IVs to ensure that the effects of the SNPs on the exposure and outcome correspond to the same alleles.

The R package "TwoSampleMR" (Bristol, UK) was used for MR analysis. Inverse variance weighting (IVW) was employed as the primary analytical method, while MR-Egger, weighted median, weighted mode, and simple mode were used as complementary methods. A sensitivity analysis was then conducted, and the second and third assumptions were indirectly tested. First, we assessed the level of horizontal pleiotropy of the IVs using the MR-PRESSO global test and MR-Egger regression. Second, Cochran's Q test was performed to evaluate the heterogeneity of IVW and MR-Egger regression. Additionally, a leave-one-out analysis was conducted to identify individual SNP that might improperly influence the MR estimates.

Statistical Analysis

The correlation analysis between the two variables was conducted using the Spearman analysis method. The Wilcoxon rank-sum test was used to analyze the differences between the two groups. The Kruskal-Wallis test was employed for differences analysis among more than two groups. K-M curves and Cox proportional hazards regression models were applied for survival analysis. Statistical analysis and graphical visualization were performed using R software (version 4.3.1) (Auckland, New Zealand). In the above analysis, $p < 0.05$ is considered statistically significant.

Results

Analysis of VPS29 in Pan-Cancer

We analyzed the mRNA expression of *VPS29* in pan-cancer using the TCGA dataset. Among the 33 types of cancer, 15 showed differential expression of *VPS29*. The expression level of *VPS29* was significantly upregulated in bladder cancer (BLCA), breast invasive carcinoma (BRCA), cervical cancer (CESC), cholangiocarcinoma (CHOL), colon adenocarcinoma (COAD), esophageal carcinoma (ESCA), glioblastoma multiforme (GBM), KIRC (kidney renal clear cell carcinoma), KIRP (kidney renal papillary cell carcinoma), LIHC (liver hepatocellular carcinoma), LUAD (lung adenocarcinoma), LUSC (lung squamous cell carcinoma), STAD (stomach adenocarcinoma), and THCA (thyroid carcinoma). Conversely, *VPS29* expression was downregulated only in KICH (kidney chromophobe) (Figure 2A). Univariate Cox regression analysis showed a significant correlation between the expression level of *VPS29* and the OS of patients in bladder cancer (BLCA), cervical cancer (CESC), esophageal carcinoma (ESCA), low-grade gliomas (LGG), liver hepatocellular carcinoma (LIHC), and adenocarcinoma of prostate (PRAD) (Figure 1B). Additionally, we validated the expression level of *VPS29* in HCC using multiple independent cohorts, showing significant upregulation of *VPS29* in HCC (Figure 2C-J).

The Prognostic and Diagnostic Value of VPS29 in HCC

K-M survival analysis revealed the prognostic role of *VPS29* in HCC. We found that increased expression of *VPS29* in HCC patients in the TCGA-LIHC dataset was significantly associated with worse OS, DSS, and PFI ($p < 0.005$) (Figure 3A-C). Through univariate and multivariate Cox regression analysis, we identified *VPS29* as a reliable and independent predictor for predicting OS in HCC patients ($p < 0.05$) (Figure 3D-E). Subsequently, a column chart was created with *VPS29* and clinical parameters to predict 1-year, 3-year, and 5-year OS in the entire cohort (Figure 3F). The ROC curve demonstrates that VPS29 exhibits a high level of accuracy in predicting and diagnosing HCC patients, particularly in terms of diagnosing HCC, with an area under the curve (AUC) of 0.925 (Figure 3G-H). To explore the correlation between *VPS29* and clinical features, we evaluated the differential expression of *VPS29* among different subgroups of clinical parameters, including age, gender, tumor stage, histological grade, tumor status, albumin (ALB), and alpha-fetoprotein (AFP) (Figure 4A-F).

Methylation Analysis

DNA methylation is a common epigenetic modification, and an increase in methylation levels often inhibits the transcription process of genes, thereby affecting gene expression. The Spearman correlation analysis results revealed a negative correlation (R = -0.23, p = 8.9e-06) between the methylation level of the *VPS29* pro-

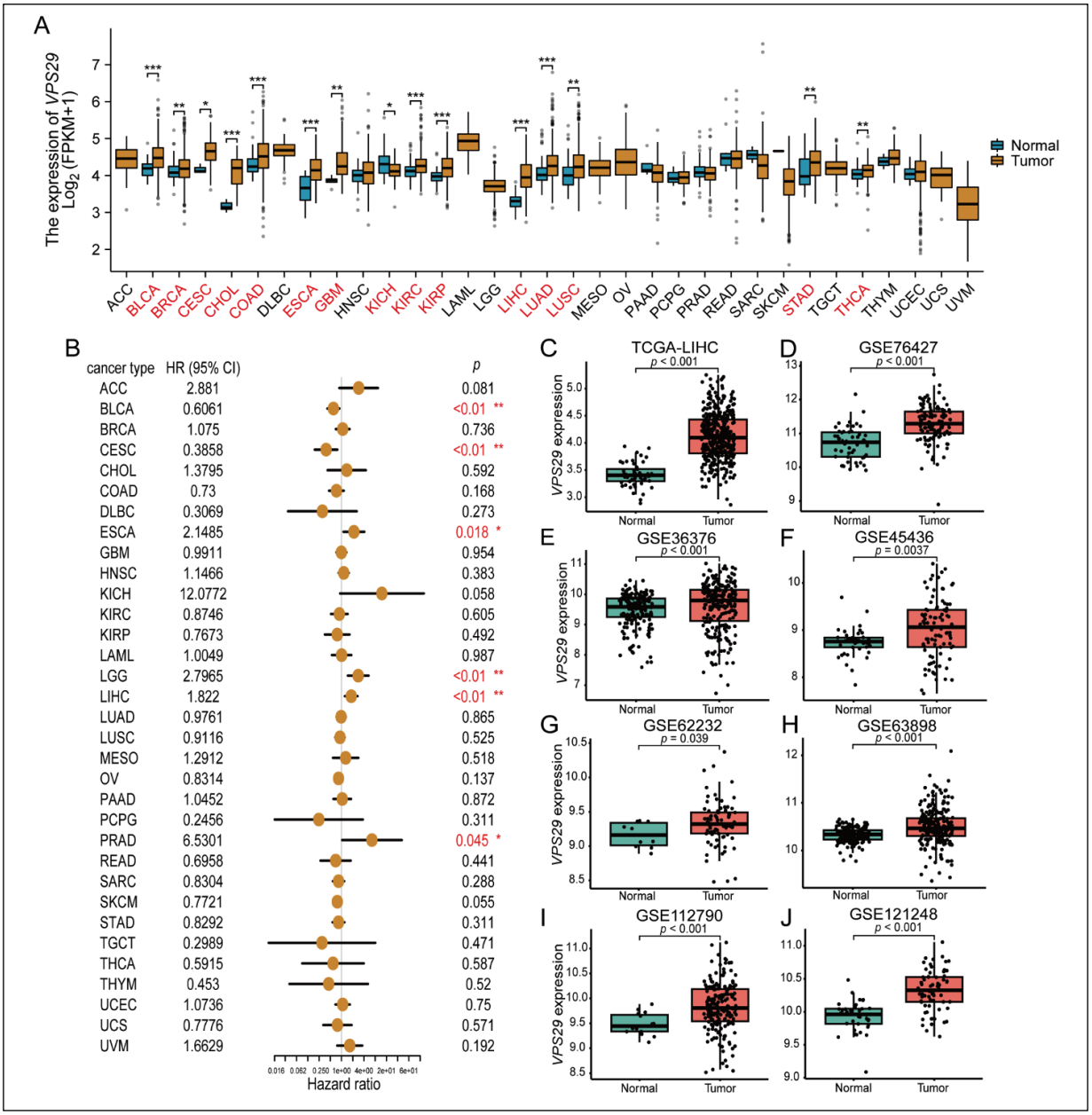

Figure 2. Analysis of *VPS29* in pan-cancer. **A**, Differential expression of *VPS29* in the TCGA pan-cancer dataset. **B**, Univariate Cox analysis of *VPS29* and OS in pan-cancer patients. **C-J**, Exploration of the expression differences of *VPS29* in HCC through multiple independent cohorts. (**p*<0.05, ***p*<0.01, ****p*<0.001).

moter region and its mRNA expression (Figure 5A). Subsequently, we identified 9 methylation sites (cg20877181, cg17382939, cg00278366, cg03867797, cg05942736, cg21605021, cg10414229, cg10025392, cg21868995), among which all sites, except cg21868995, exhibited a negative correlation with *VPS29* expression (Figure 5B-J). Based on the median methylation level at each site, HCC patients were divided into high methylation and low methylation groups, and K-M survival analysis was performed to compare OS between the two groups. The results indicated that high methylation levels at the cg20877181, cg03867797, cg10025392, and cg21605021 sites were associated with better prognosis (Figure 5K-N). These findings suggest that the methylation level of *VPS29* may serve as a potential prognostic biomarker and may play an important role in the progression of HCC.

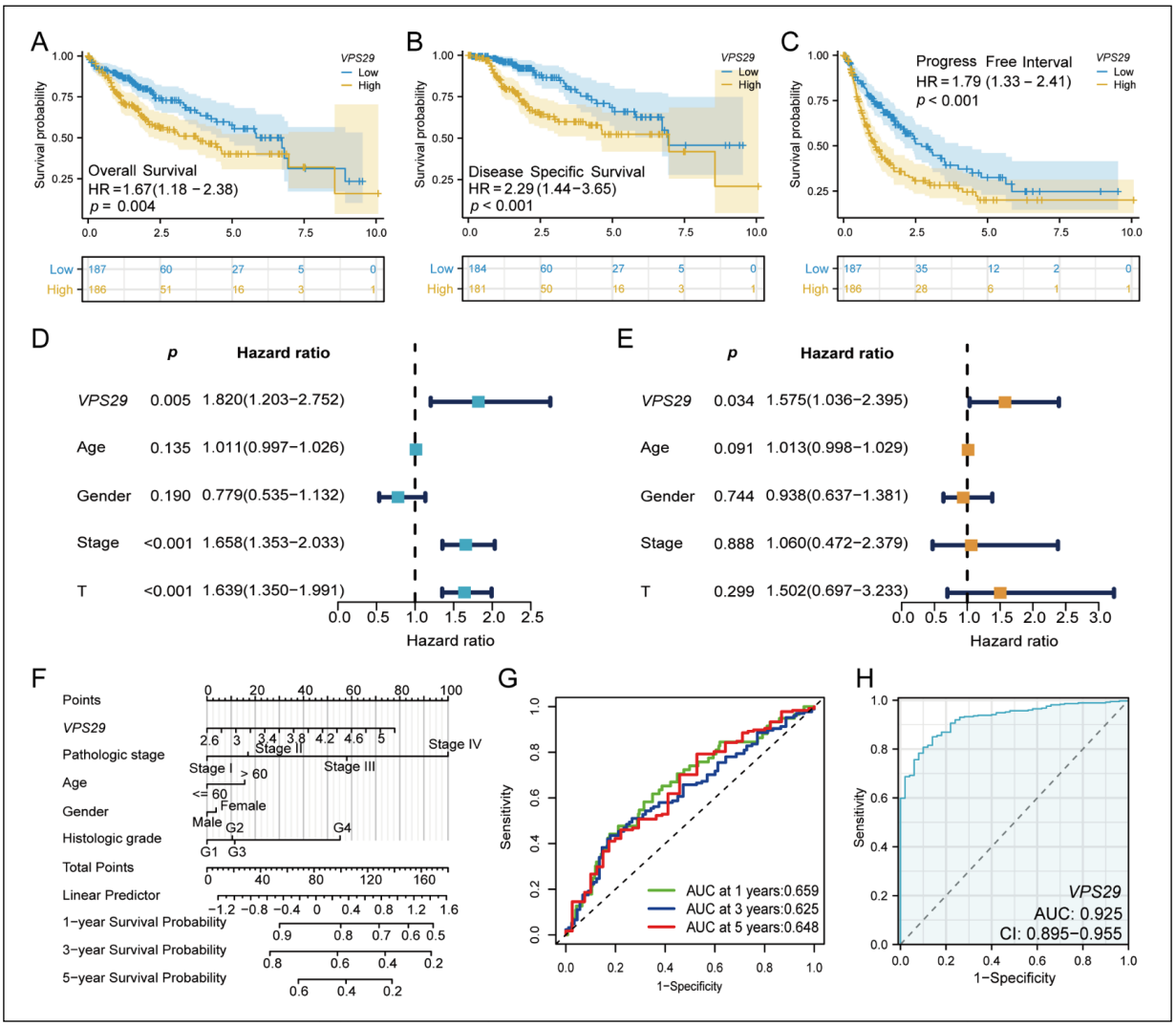

Figure 3. Prognostic and diagnostic value of *VPS29* in HCC. **A-C**, K-M survival curves for OS, DSS, and PFI between high and low *VPS29* expression group. **D-E**, Univariate and multivariate regression analyses between clinical characteristics and OS. **F**, Nomogram for predicting 1 - year, 3 - year, and 5-year OS in HCC patients. **G**, Time-dependent ROC curve for predicting OS by *VPS29*. **H**, ROC curve for *VPS29* in diagnosing HCC.

Drug Sensitivity Analysis

Systemic drug treatment is crucial for HCC, and to explore whether *VPS29* can be used for predicting drug response in HCC patients, we compared the differences in IC50 values for 15 commonly used antineoplastic drugs between the high and low *VPS29* expression groups. The results revealed that in the high *VPS29* expression group, six drugs showed lower IC50 values (indicating better drug sensitivity), including Cediranib (Cambridge, London, UK), Osimertinib (Cambridge, London, UK), Docetaxel (Paris, France), Paclitaxel (Bristol, Pennsylvania, USA), Vinblastine (Madison, Wisconsin, USA) and Vinorelbine (Paris, France), suggesting potential benefits for these patients (Figure 6A-F). Additionally, high *VPS29* expression patients may exhibit resistance to four drugs, including Gefitinib (Wilmington, Delaware, USA), Mitoxantrone (Lawrenceville, New Jersey, USA), Oxaliplatin (Paris, France), and Sorafenib (Leverkusen, North Rhine-Westphalia, Germany) (Figure 6G-J). The remaining five drugs showed no correlation with *VPS29* (Figure 6K-O).

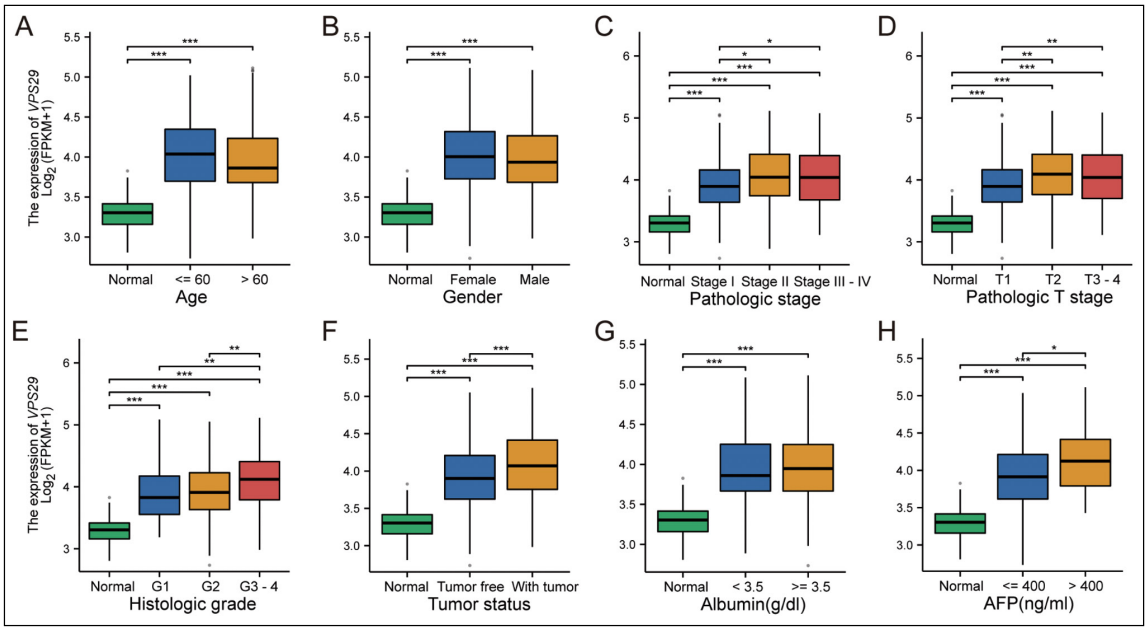

Figure 4. Correlation between *VPS29* expression levels and clinical information. **A,** Age. **B,** Gender. **C,** Pathologic stage. **D,** T stage. **E,** Histologic grade. **F,** Tumor status. **G,** ALB. **H,** AFP. (**p*<0.05, ***p*<0.01, ****p*<0.001).

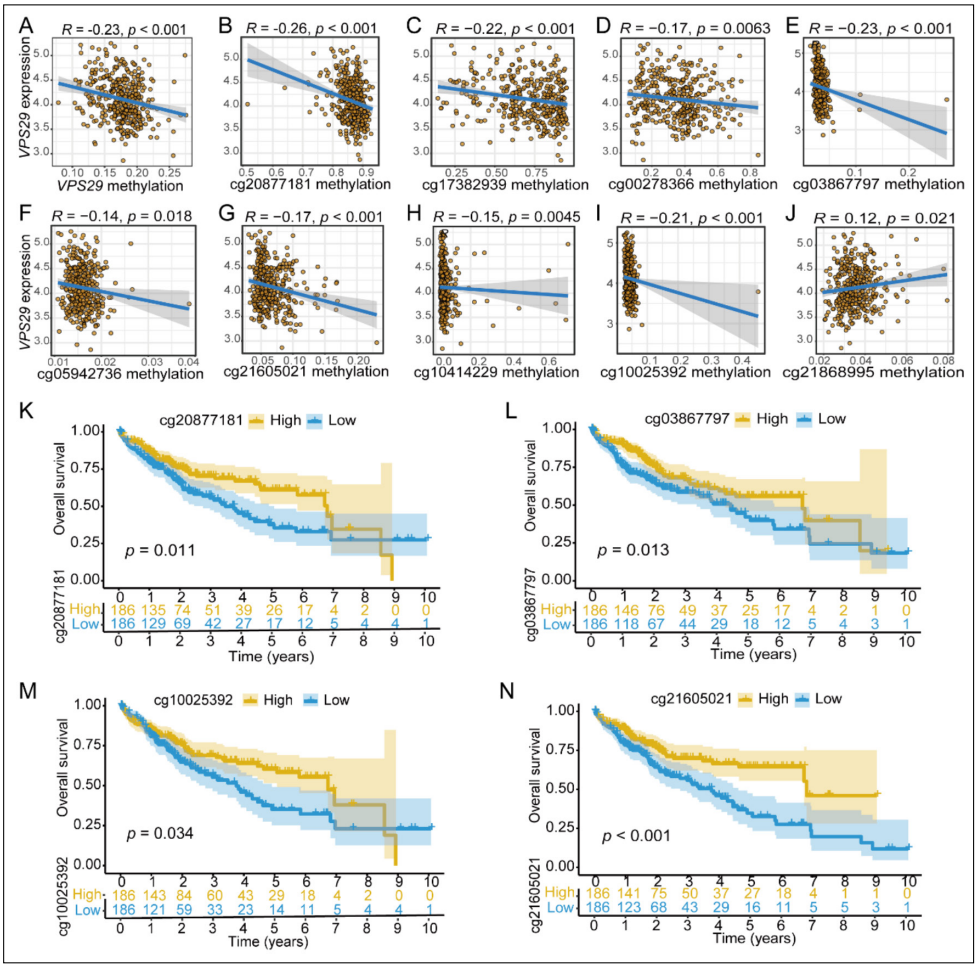

Figure 5. DNA methylation analysis of *VPS29* in HCC. **A**, Correlation between *VPS29* methylation and its mRNA expression. **B-J**, Correlation between *VPS29* and methylation sites. **K-N**, KM survival analysis based on methylation sites.

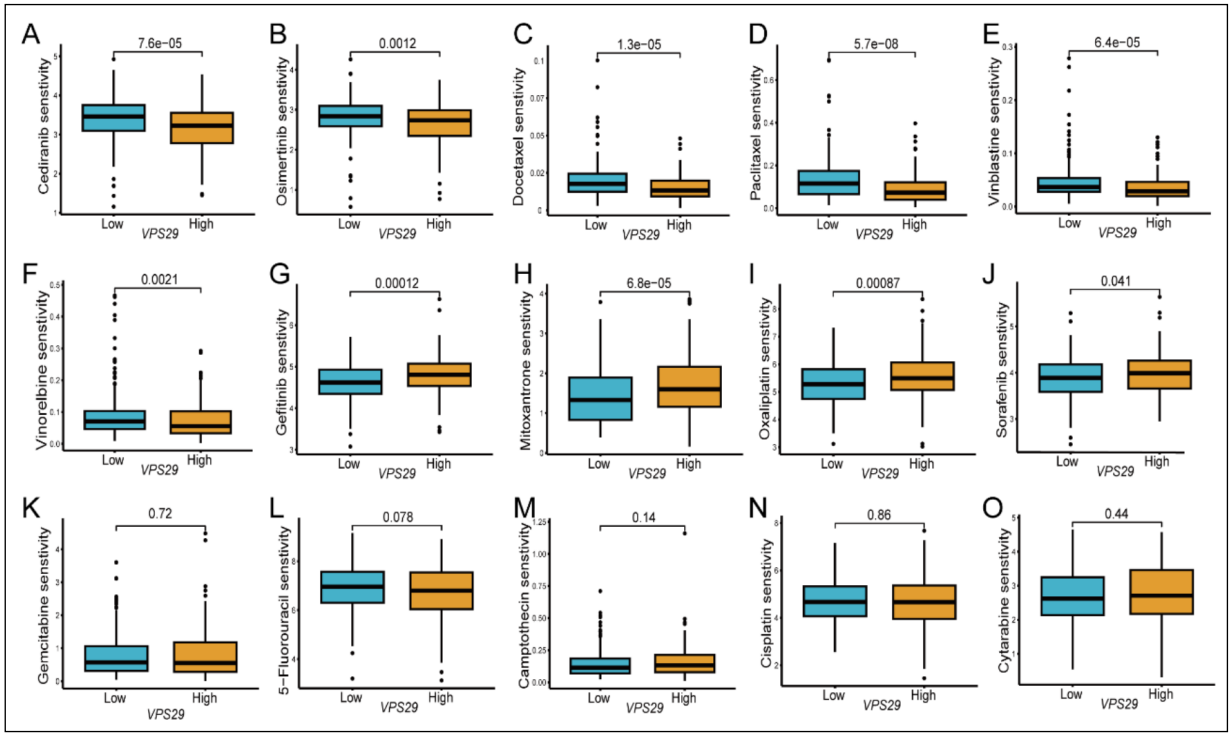

Figure 6. Drug sensitivity analysis. **A**, Cediranib. **B**, Osimertinib. **C**, Docetaxel. **D**, Paclitaxel. **E**, Vinblastine. **F**, Vinorelbine. **G**, Gefitinib. **H**, Mitoxantrone. **I**, Oxaliplatin. **J**, Sorafenib. **K**, Gemcitabine. **L**, 5-Fluorouracil. **M**, Camptothecin. **N**, Cisplatin. **O**, Cytarabine.

Correlation Analysis of VPS29 with Tumor Immune Microenvironment and Tumor Mutation Burden

We compared immune cell enrichment scores between the high and low *VPS29* expression groups to determine the correlation between *VPS29* expression and the immune infiltration level in HCC. The results showed that NK CD-56bright cells, TFH, and Th2 cells had higher levels of infiltration in the high *VPS29* expression group, while CD8 T cells, Eosinophils, Neutrophils, Tcm, and Th17 cells had higher levels in the low *VPS29* expression group (Figure 7A). Additionally, the expression level of *VPS29* was significantly correlated with the infiltration levels of various immune cells, such as Th2 cells, T helper cells, Th17 cells, Neutrophils, and Eosinophils ($p < 0.05$) (Figure 7B). To further evaluate the correlation between *VPS29* and immune therapy response, we calculated TIDE and TMB scores for HCC patients. The results showed that the TIDE score was significantly higher in the high VPS29 expression group, indicating a poorer immune therapy response in patients with high *VPS29* expression (Figure 7C). However, no significant correlation was observed between TMB and *VPS29* (Figure 7D). In summary, these results suggest that *VPS29* may be closely associated with immune infiltration in the progression of HCC.

Identification of Co-Expressed Genes and Differentially Expressed Genes

We identified 10 significantly correlated co-expressed genes with *VPS29* through Spearman analysis ($|R| > 0.6$, $p < 0.001$), all of which showed a positive correlation with *VPS29* (Figure 8A). 9 of these genes were associated with poor prognosis in HCC, including *ARPC3*, *RFC5*, *CCDC59*, *RFC4*, *MAPKAPK5*, *ANAPC7*, *BRAP*, *PWP1*, and *ANAPC5* (Figure 8B). Subsequently, we performed differential expression analysis on the high and low *VPS29* expression groups and identified 602 DEGs, including 573 upregulated genes and 29 downregulated genes in HCC. The heatmap depicts the expression profiles of the top 50 DEGs (Figure 9A).

To further understand the underlying molecular biological mechanisms between the high and low *VPS29* expression groups, we conducted an enrichment analysis on the 602 differentially expressed genes. The top 30 enriched terms in the

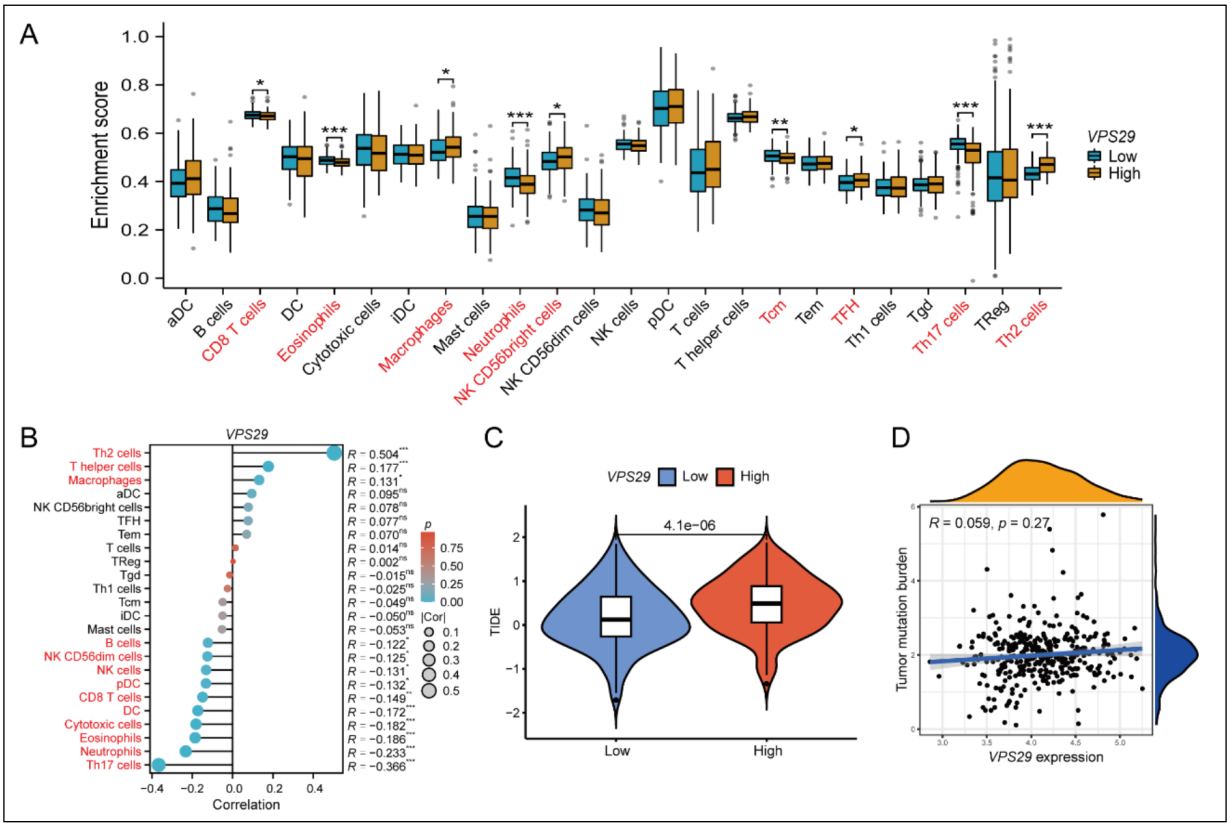

Figure 7. Correlation analysis of *VPS29* expression with the TIME in HCC. **A**, Differences in immune cell infiltration between the high and low *VPS29* expression groups. **B**, Correlation analysis between *VPS29* and immune cell infiltration levels. **C**, Differences in TIDE scores between the high and low *VPS29* expression groups. **D**, Correlation analysis between *VPS29* and TMB scores. (**p*<0.05, ***p*<0.01, ****p*<0.001).

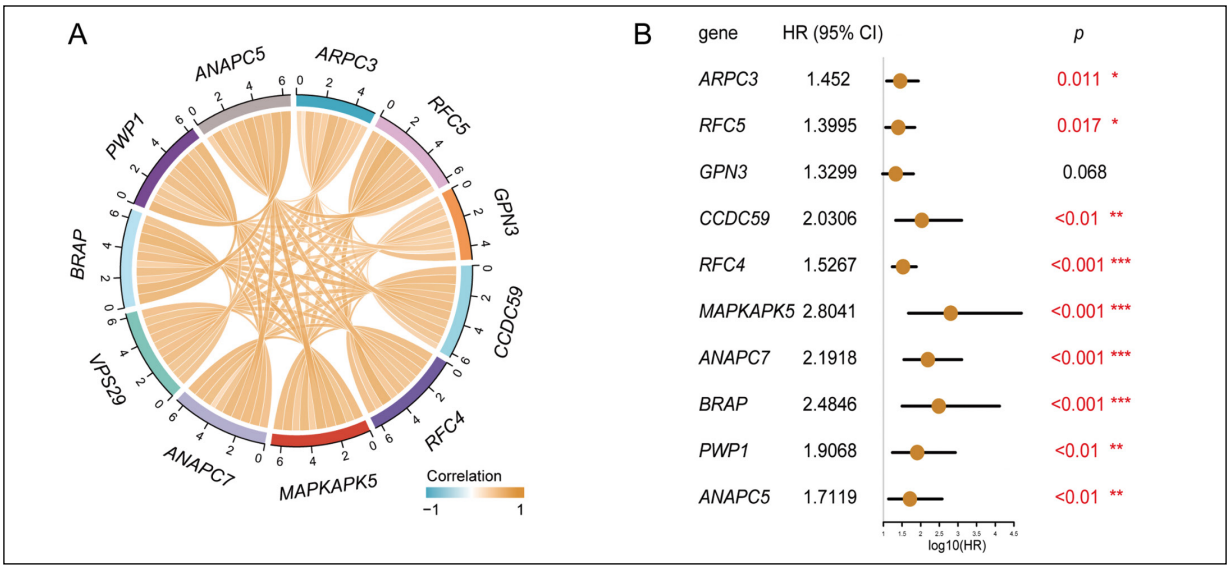

Figure 8. Analysis of co-expressed genes. **A**, Top 10 co-expressed genes most significantly correlated with *VPS29*. **B**, Univariate Cox analysis of co-expressed genes.

GO analysis were identified as the main functions (Figure 9B), including meiotic cell cycle and nuclear division in biological processes (BP), kinesin complex and P granule in cellular components

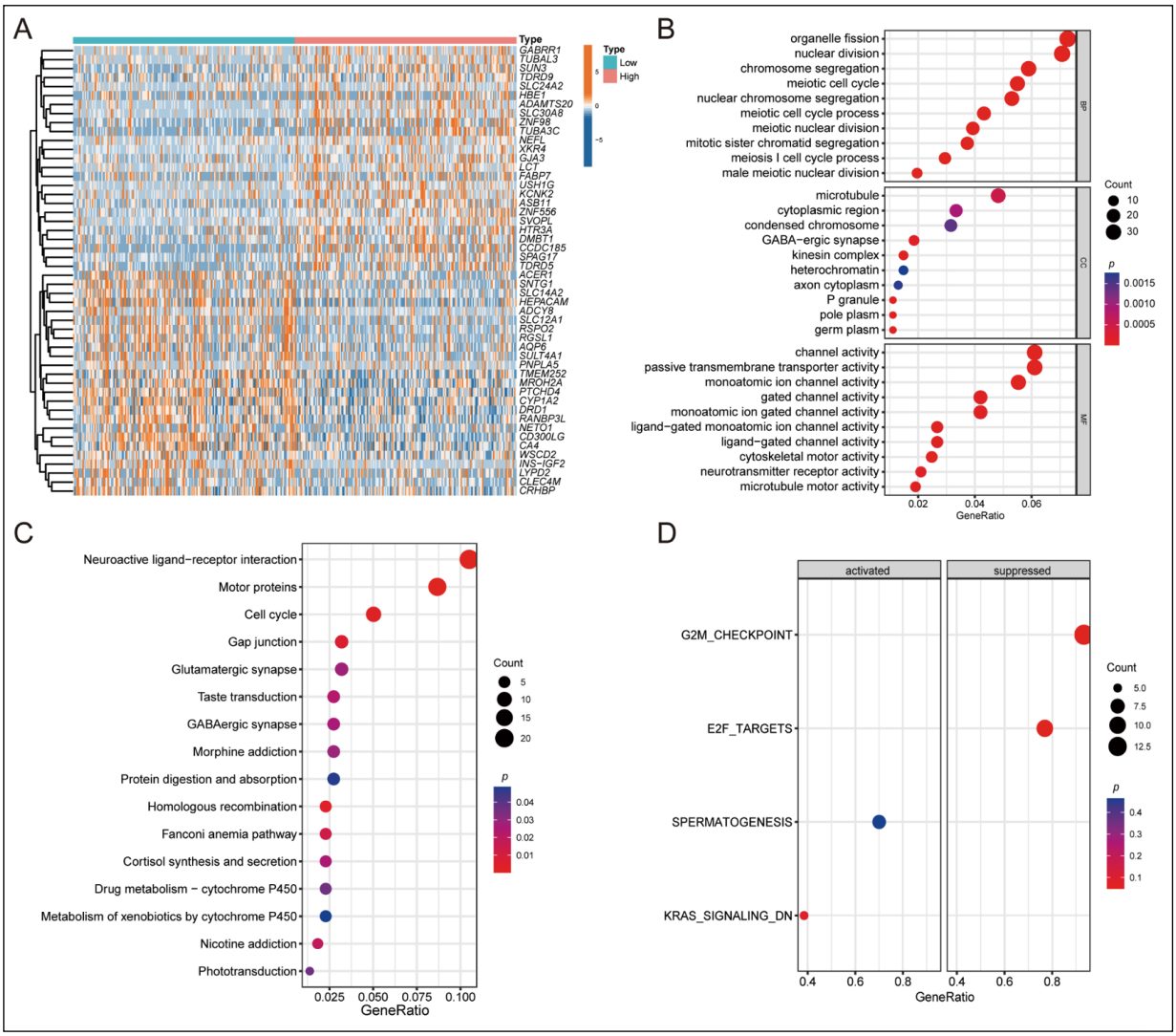

Figure 9. Analysis of differentially expressed genes. **A**, Heatmap of differentially expressed genes. **B**, GO enrichment analysis. **C**, KEGG enrichment analysis. **D**, GSEA enrichment analysis.

(CC), and channel activity and passive transmembrane transporter activity in molecular functions (MF). KEGG enrichment identified a total of 16 signaling pathways, mainly including Cell cycle, Homologous recombination, Gap junction, and Neuroactive ligand-receptor interaction (Figure 9C). In the analysis of GSEA, 4 major pathways were identified (Figure 9D).

Two-Sample Mendelian Randomization Analysis

After applying 5 filtering criteria, we obtained 19 SNP as IVs that met the criteria. Detailed information on these SNPs can be found in **Supplementary Table I**. We employed 5 MR analysis methods to evaluate the causal effect between VPS29 and HCC. The IVW analysis revealed that VPS29 protein is a risk factor for HCC (OR: 1.440; 95% CI: 1.195-1.736). The other 3 analysis methods also yielded consistent results with IVW, including MR Egger (OR: 1.553; 95% CI: 1.220-1.977), Weighted median (OR: 1.462; 95% CI: 1.184-1.805), and Weighted mode (OR: 1.472; 95% CI: 1.190-1.822), further validating the results of IVW analysis (Figure 10A). We also assessed the impact of each SNP on HCC (Figure 10B). As depicted in the scatter plot, the direction of all 5 MR analysis methods was consistent, indicating that the risk of developing HCC increases with increasing levels of VPS29 (Figure 10C).

Subsequently, we conducted sensitivity analyses to evaluate the robustness of our results. Firstly, the results of Cochran's Q test showed

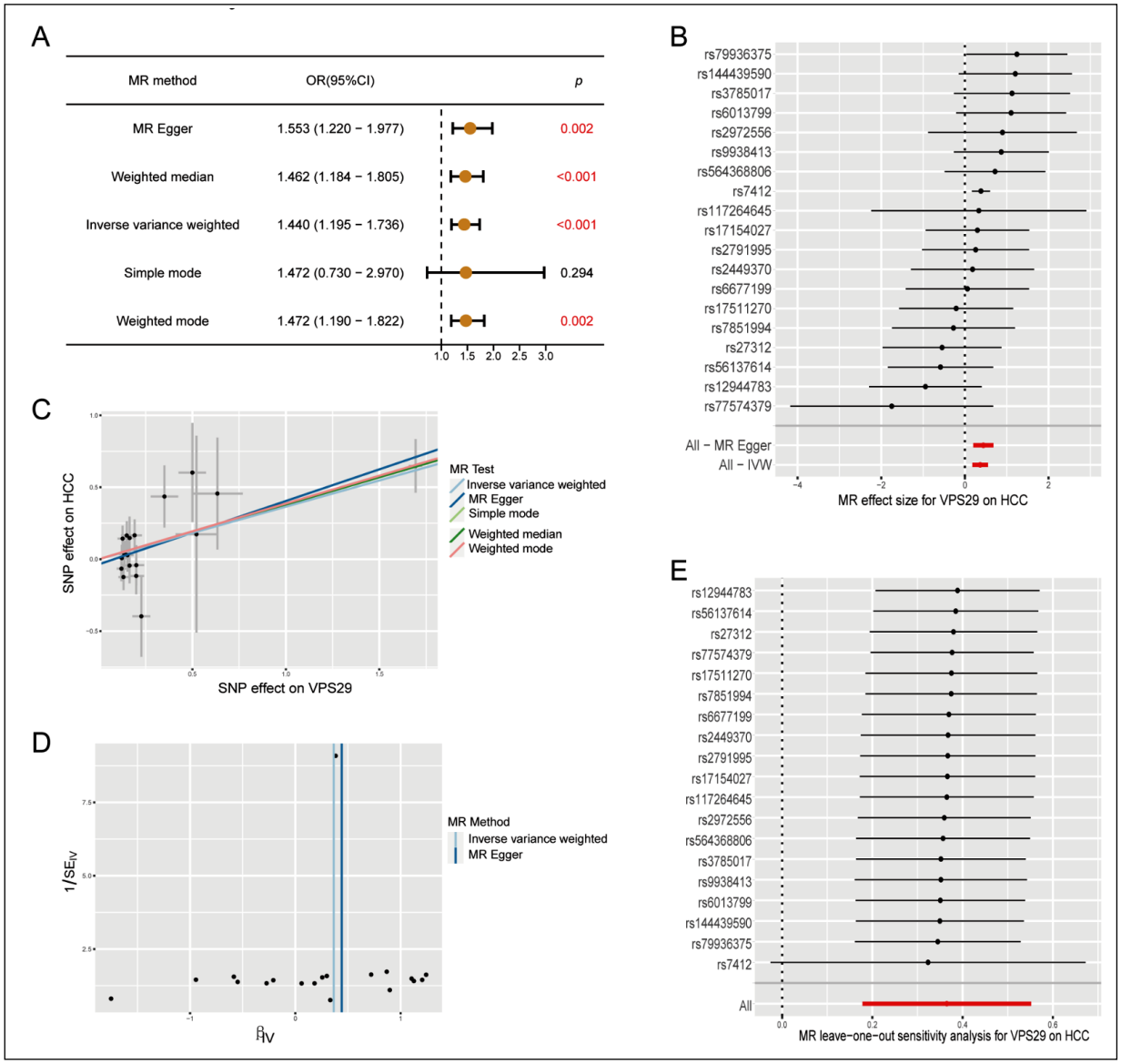

Figure 10. MR analysis of the causal relationship between VPS29 protein and HCC. **A**, Forest plot of the 5 MR analysis methods. **B**, Forest plot of the causal effects of each SNP on HCC risk. **C**, Scatter plot of the 5 MR analysis methods. **D**, Funnel plot for heterogeneity assessment. **E**, Sensitivity analysis using the leave-one-out method.

no heterogeneity among the IVs (p_{IVW} = 0.371, $p_{MR\ Egger}$ = 0.368) (**Supplementary Table II**). The symmetry of the funnel plot also confirmed the absence of heterogeneity (Figure 10D). Secondly, the MR-PRESSO global test indicated the absence of horizontal pleiotropy across all IVs (p = 0.511), and the MR Egger regression test ensured the accuracy of the results (p = 0.346) (**Supplementary Table II**). The leave-one-out analysis did not observe significant changes in the estimated causal effects when excluding each SNP, further supporting the reliability of the MR analysis results (Figure 10E).

Discussion

The liver is a vital organ involved in diverse physiological processes, such as the metabolism of macromolecules, control of endocrine growth signals, stabilization of lipid and cholesterol levels, and regulation of other physiological reactions[17]. Therefore, the problems associated with the liver should not be underestimated. HCC is one of the most common malignancies worldwide, and most patients are diagnosed at an advanced stage, often accompanied by lower quality of life and poor prognosis[18]. Therefore,

it is still essential to identify new promising targets for the diagnosis and treatment of HCC. The highly conserved retromer complex controls the fate of hundreds of receptors through the endolysosomal system and acts as a central regulatory node in various metabolic processes[19]. Studies[20] have found the involvement of the retromer complex in the occurrence and progression of cancer. A common example is VPS35, which binds the farnesylated form of the oncoprotein N-RAS and affects its trafficking and membrane association[21]. Frameshift mutations and low expression of *VPS35* have been detected in over half of analyzed cancer samples[14]. VPS29 protein, as a core subunit of the retrograde complex, has been poorly studied in cancer, including HCC. Therefore, this study used comprehensive bioinformatics techniques combined with MR analysis methods to investigate the expression, biological functions, prognostic role, and diagnostic value of *VPS29* in HCC, as well as evaluate the causal relationship between VPS29 protein and HCC.

In this study, we analyzed the TCGA database and found that *VPS29* was significantly overexpressed in multiple cancer types, including HCC, and its dysregulation was associated with patient prognosis. We validated the significant overexpression of *VPS29* in HCC using multiple independent cohorts. Further analysis revealed that increased *VPS29* expression in HCC patients was significantly associated with poorer OS, DSS, and PFI. *VPS29* was identified as an independent predictor of OS in HCC patients ($p < 0.05$). ROC curve analysis demonstrated that *VPS29* was able to discriminate between HCC tissues and normal tissues effectively. Differential expression of *VPS29* was identified in various clinical subgroup analyses, including tumor stage, T stage, histological grade, tumor status, and AFP levels. We also identified the top 10 genes most significantly correlated with *VPS29* in HCC tissues, all of which showed a positive correlation with *VPS29*. Among these genes, 9 were significantly associated with poor prognosis in HCC, suggesting their potential cooperation in promoting HCC progression. Therefore, *VPS29* may serve as a molecular biomarker for the diagnosis and prognosis of HCC patients.

Abnormal DNA methylation has been recognized as a common epigenetic alteration in human cancer[22]. Therefore, DNA methylation has received increasing attention in tumor occurrence research, early diagnosis, and prognostic prediction. Studies have shown that DNA methylation plays an important role in tumor initiation and progression and is considered one of the key mechanisms in tumor development[23,24]. In HCC, the detection of DNA methylation markers has been found to be specific[25]. Our study found a negative correlation between *VPS29* expression and methylation levels, identifying 9 methylation sites in the promoter region of *VPS29* that were significantly correlated, with 8 of them showing a negative correlation with *VPS29* expression level. Further survival analysis revealed that low methylation levels at the cg20877181, cg03867797, cg10025392, and cg21605021 sites were associated with worse OS. Therefore, we speculate that the loss or reduction of methylation promotes *VPS29* expression, leading to the occurrence and progression of HCC.

Advanced HCC generally lacks surgical indications, and patients can only hope to benefit from systemic drug therapy. Paradoxically, researchers have reached a consensus that the later the stage of HCC, the stronger its heterogeneity[26]. Currently, commonly used treatments, such as sorafenib, are effective in less than 40% of HCC patients[27]. Therefore, there is an urgent need for biomarkers to predict drug responses. We evaluated the potential correlation between *VPS29* and the IC50 values of commonly used antineoplastic drugs. The results showed that Cediranib, Osimertinib, Docetaxel, Paclitaxel, Vinblastine, and Vinorelbine had better responses in patients with high *VPS29* expression, while Gefitinib, Mitoxantrone, Oxaliplatin, and Sorafenib were more sensitive in patients with low *VPS29* expression. These findings suggest that *VPS29* can indicate the drug response in different expression subgroups of HCC patients.

The tumor microenvironment (TME) is a highly complex structure mainly composed of tumor cells, stromal cells, immune cells, cancer-associated fibroblasts (CAFs), blood vessels, and various signaling molecules[28,29]. Among them, the tumor immune microenvironment (TIME) forms the basis for the interaction between tumor cells and immune cells, which is crucial for the progression of HCC and has a significant impact on the outcomes of immune therapies[30]. Therefore, the identification of new predictive biomarkers and a comprehensive understanding of the immune infiltration status in cancer patients are important for selecting accurate, personalized immune therapies. Our study found that NK CD56 bright cells, TFH, and Th2 cells had higher levels of infiltration in the high *VPS29* expression group, while CD8 T cells, Eosinophils, Neutrophils, Tcm, and Th17 cells had higher levels of infiltration in the low *VPS29* expression group. Among

them, Th2 cells showed the strongest correlation with *VPS29*. Th2 cell infiltration has been associated with immune suppression and lower survival rates in many malignancies[31]. Th2 cells can inhibit the development of Th1 cells and the release of IFN-γ, while also secreting IL-4 and IL-10, thereby promoting tumor cell growth[32]. Multiple studies[33,34] have indicated the important role of T cells in the immune defense against HCC. The occurrence of liver diseases, including liver tumors, severely impacts the immune function of the liver and has a significant influence on the progression of liver cancer[35]. Lurje et al[36] found that by mediating tolerance or immunogenicity of T cell responses and activating CD8+ T cells, effective anti-tumor defense mechanisms can be established in the liver. Neutrophils are innate immune cells that infiltrate tissues during infections, injuries, or cancer. They serve as important effector cells in anti-cancer immunity. Neutrophils account for a significant portion of circulating white blood cells, and tumor-associated neutrophils (TAN) produce chemotactic factors and pro-inflammatory cytokines that recruit and activate CD8+ T cells, thereby limiting tumor growth[37]. HCC typically exhibits a "cold" immune state, protecting cancer cells from the attack of infiltrating lymphocytes, which may result in poor response to immune therapy[38]. Our study also found that HCC patients with high expression of *VPS29* had higher TIDE scores, indicating a poorer response to immune therapy in patients with increased *VPS29* expression. These findings suggest that *VPS29* plays an indispensable role in regulating immune cell infiltration and the efficacy of immune therapy in HCC.

Finally, we conducted a two-sample MR analysis to explore the causal relationship between VPS29 protein and HCC. The results showed that VPS29 protein significantly increases the risk of HCC. The IVW method, along with the remaining complementary methods, consistently yielded OR values in the same direction, with 4 methods providing strong evidence ($p < 0.01$). Sensitivity analysis indicated that our results were robust.

Limitations

In conclusion, our study provides new insights and evidence for personalized clinical treatment strategies for HCC patients, although it should be acknowledged that there are some limitations to this research. Firstly, our study only involved data mining from publicly available databases and lacked validation on physical clinical specimens. Additionally, the effect estimates in the MR analysis were focused on lifelong exposure to genetic susceptibility, without considering the temporal and spatial fluctuations in gene expression within tissues. However, this study validated the differential expression of *VPS29* in HCC across multiple independent cohorts, and the MR analysis provided robust evidence for this research. Therefore, the results are still acceptable. Future research involving prospective clinical trials and mechanistic exploration is necessary to validate the current findings further.

Conclusions

Our study indicates that the *VPS29* gene and VPS29 protein significantly increase the risk of HCC. Its upregulation in HCC is associated with poor prognosis and is correlated with methylation levels, immune microenvironment, and drug sensitivity in HCC. These findings suggest that VPS29 may serve as a biomarker and potential therapeutic target for HCC.

Conflict of Interest

The authors declare that they have no conflict of interests.

Funding

This work was supported by The National Natural Science Foundation of China (Grant No. 81860839) and the Natural Science Foundation of Guangxi (Grant 2022GXNSFAA035446).

Authors' Contribution

Authors Yuanqian Yao and Jianlin Lv have made the same contribution to this paper, which includes concept proposing, data management, formal analysis, survey development, research method development or design, management, and coordination of research project planning and execution, and resource provision. Both of them supervised and led the whole research activity planning and execution.

Ethics Approval

Not applicable.

Availability of Data and Materials

The datasets used and analyzed during the current study are available from The Cancer Genome Atlas (TCGA, https://portal.gdc.cancer.gov/), Gene expression Omnibus (GEO, https://www.ncbi.nlm.nih.gov/geo/), and integrative epidemiology unit Open genome-wide association studies database (IEU Open GWAS, https://gwas.mrcieu.ac.uk/).

Informed Consent
Not applicable.

ORCID ID
Yuanqian Yao: 0000-0002-8175-0772
Jianlin Lv: 0000-0002-3237-4301

References

1) Yang Y, Xiong L, Li M, Jiang P, Wang J, Li C. Advances in radiotherapy and immunity in hepatocellular carcinoma. J Transl Med 2023; 21: 526.
2) McGlynn KA, Petrick JL, El-Serag HB. Epidemiology of Hepatocellular Carcinoma. Hepatology 2021; 73 Suppl 1: 4-13.
3) Villanueva A. Hepatocellular Carcinoma. N Engl J Med 2019; 380: 1450-1462.
4) Minaei N, Ramezankhani R, Tamimi A, Piryaei A, Zarrabi A, Aref AR, Mostafavi E, Vosough M. Immunotherapeutic approaches in Hepatocellular carcinoma: Building blocks of hope in near future. Eur J Cell Biol 2023; 102: 151284.
5) Craig AJ, von Felden J, Garcia-Lezana T, Sarcognato S, Villanueva A. Tumour evolution in hepatocellular carcinoma. Nat Rev Gastroenterol Hepatol 2020; 17: 139-152.
6) Baños-Mateos S, Rojas AL, Hierro A. VPS29, a tweak tool of endosomal recycling. Curr Opin Cell Biol 2019; 59: 81-87.
7) McNally KE, Faulkner R, Steinberg F, Gallon M, Ghai R, Pim D, Langton P, Pearson N, Danson CM, Nägele H, Morris LL, Singla A, Overlee BL, Heesom KJ, Sessions R, Banks L, Collins BM, Berger I, Billadeau DD, Burstein E, Cullen PJ. Retriever is a multiprotein complex for retromer-independent endosomal cargo recycling. Nat Cell Biol 2017; 19: 1214-1225.
8) Singla A, Fedoseienko A, Giridharan S, Overlee BL, Lopez A, Jia D, Song J, Huff-Hardy K, Weisman L, Burstein E, Billadeau DD. Endosomal PI(3)P regulation by the COMMD/CCDC22/CCDC93 (CCC) complex controls membrane protein recycling. Nat Commun 2019; 10: 4271.
9) Chmiest D, Sharma N, Zanin N, Viaris de Lesegno C, Shafaq-Zadah M, Sibut V, Dingli F, Hupé P, Wilmes S, Piehler J, Loew D, Johannes L, Schreiber G, Lamaze C. Spatiotemporal control of interferon-induced JAK/STAT signalling and gene transcription by the retromer complex. Nat Commun 2016; 7: 13476.
10) Li B, Wong C, Gao SM, Zhang R, Sun R, Li Y, Song Y. The retromer complex safeguards against neural progenitor-derived tumorigenesis by regulating Notch receptor trafficking. Elife 2018; 7: e38181.
11) McGough IJ, de Groot R, Jellett AP, Betist MC, Varandas KC, Danson CM, Heesom KJ, Korswagen HC, Cullen PJ. SNX3-retromer requires an evolutionary conserved MON2:DOPEY2:ATP9A complex to mediate Wntless sorting and Wnt secretion. Nat Commun 2018; 9: 3737.
12) Trousdale C, Kim K. Retromer: Structure, function, and roles in mammalian disease. Eur J Cell Biol 2015; 94: 513-521.
13) Zhang G, Tang X, Liang L, Zhang W, Li D, Li X, Zhao D, Zheng Y, Chen Y, Hao B, Wang K, Tang N, Ding K. DNA and RNA sequencing identified a novel oncogene VPS35 in liver hepatocellular carcinoma. Oncogene 2020; 39: 3229-3244.
14) An CH, Kim YR, Kim HS, Kim SS, Yoo NJ, Lee SH. Frameshift mutations of vacuolar protein sorting genes in gastric and colorectal cancers with microsatellite instability. Hum Pathol 2012; 43: 40-47.
15) Richmond RC, Davey Smith G. Mendelian Randomization: Concepts and Scope. Cold Spring Harb Perspect Med 2022; 12: a040501.
16) Davey Smith G, Hemani G. Mendelian randomization: genetic anchors for causal inference in epidemiological studies. Hum Mol Genet 2014; 23: R89-R98.
17) Trefts E, Gannon M, Wasserman DH. The liver. Curr Biol 2017; 27: R1147-R1151.
18) Miller KD, Nogueira L, Devasia T, Mariotto AB, Yabroff KR, Jemal A, Kramer J, Siegel RL. Cancer treatment and survivorship statistics, 2022. CA Cancer J Clin 2022; 72: 409-436.
19) Carosi JM, Denton D, Kumar S, Sargeant TJ. Receptor Recycling by Retromer. Mol Cell Biol 2023; 43: 317-334.
20) Yong C, Tang BL. Cancer-driving mutations and variants of components of the membrane trafficking core machinery. Life Sci 2021; 264: 118662.
21) Zhou M, Wiener H, Su W, Zhou Y, Liot C, Ahearn I, Hancock JF, Philips MR. VPS35 binds farnesylated N-Ras in the cytosol to regulate N-Ras trafficking. J Cell Biol 2016; 214: 445-458.
22) Mehdi A, Rabbani SA. Role of Methylation in Pro- and Anti-Cancer Immunity. Cancers (Basel) 2021; 13: 545.
23) Horie M, Kaczkowski B, Ohshima M, Matsuzaki H, Noguchi S, Mikami Y, Lizio M, Itoh M, Kawaji H, Lassmann T, Carninci P, Hayashizaki Y, Forrest A, Takai D, Yamaguchi Y, Micke P, Saito A, Nagase T. Integrative CAGE and DNA Methylation Profiling Identify Epigenetically Regulated Genes in NSCLC. Mol Cancer Res 2017; 15: 1354-1365.
24) Zhu Y, Lu H, Zhang D, Li M, Sun X, Wan L, Yu D, Tian Y, Jin H, Lin A, Gao F, Lai M. Integrated analyses of multi-omics reveal global patterns of methylation and hydroxymethylation and screen the tumor suppressive roles of HADHB in colorectal cancer. Clin Epigenetics 2018; 10: 30.
25) Fu S, Debes JD, Boonstra A. DNA methylation markers in the detection of hepatocellular carcinoma. Eur J Cancer 2023; 191: 112960.

26) Vij M, Calderaro J. Pathologic and molecular features of hepatocellular carcinoma: An update. World J Hepatol 2021; 13: 393-410.

27) Sperandio RC, Pestana RC, Miyamura BV, Kaseb AO. Hepatocellular Carcinoma Immunotherapy. Annu Rev Med 2022; 73: 267-278.

28) Binnewies M, Roberts EW, Kersten K, Chan V, Fearon DF, Merad M, Coussens LM, Gabrilovich DI, Ostrand-Rosenberg S, Hedrick CC, Vonderheide RH, Pittet MJ, Jain RK, Zou W, Howcroft TK, Woodhouse EC, Weinberg RA, Krummel MF. Understanding the tumor immune microenvironment (TIME) for effective therapy. Nat Med 2018; 24: 541-550.

29) Hinshaw DC, Shevde LA. The Tumor Microenvironment Innately Modulates Cancer Progression. Cancer Res 2019; 79: 4557-4566.

30) Kotsari M, Dimopoulou V, Koskinas J, Armakolas A. Immune System and Hepatocellular Carcinoma (HCC): New Insights into HCC Progression. Int J Mol Sci 2023; 24: 11471.

31) Schreiber S, Hammers CM, Kaasch AJ, Schraven B, Dudeck A, Kahlfuss S. Metabolic Interdependency of Th2 Cell-Mediated Type 2 Immunity and the Tumor Microenvironment. Front Immunol 2021; 12: 632581.

32) Lee HL, Jang JW, Lee SW, Yoo SH, Kwon JH, Nam SW, Bae SH, Choi JY, Han NI, Yoon SK. Inflammatory cytokines and change of Th1/Th2 balance as prognostic indicators for hepatocellular carcinoma in patients treated with transarterial chemoembolization. Sci Rep 2019; 9: 3260.

33) Woller N, Engelskircher SA, Wirth T, Wedemeyer H. Prospects and Challenges for T Cell-Based Therapies of HCC. Cells 2021; 10: 1651.

34) Zheng C, Zheng L, Yoo JK, Guo H, Zhang Y, Guo X, Kang B, Hu R, Huang JY, Zhang Q, Liu Z, Dong M, Hu X, Ouyang W, Peng J, Zhang Z. Landscape of Infiltrating T Cells in Liver Cancer Revealed by Single-Cell Sequencing. Cell 2017; 169: 1342-1356.e16.

35) Chen Y, Tian Z. HBV-Induced Immune Imbalance in the Development of HCC. Front Immunol 2019; 10: 2048.

36) Lurje I, Hammerich L, Tacke F. Dendritic Cell and T Cell Crosstalk in Liver Fibrogenesis and Hepatocarcinogenesis: Implications for Prevention and Therapy of Liver Cancer. Int J Mol Sci 2020; 21: 7378.

37) Sionov RV, Fridlender ZG, Granot Z. The Multifaceted Roles Neutrophils Play in the Tumor Microenvironment. Cancer Microenviron 2015; 8: 125-158.

38) Wang Y, Zhao Q, Zhao B, Zheng Y, Zhuang Q, Liao N, Wang P, Cai Z, Zhang D, Zeng Y, Liu X. Remodeling Tumor-Associated Neutrophils to Enhance Dendritic Cell-Based HCC Neoantigen Nano-Vaccine Efficiency. Adv Sci (Weinh) 2022; 9: e2105631.