

# High expression of HILPDA is an adverse prognostic prognostic factor in hepatocellular carcinoma

Xiao Wang, BM<sup>a</sup> , Aoshuang Zou, MM<sup>b</sup>, Jinhe Zhang, BM<sup>a</sup>, Guochuan Gao, MM<sup>b</sup>, Wenting Shan, BM<sup>a</sup>, Jun Li, MD<sup>a</sup>, Xia Liu, MD<sup>a,\*</sup>

## Abstract

Hepatocellular carcinoma (LIHC) is a malignant tumor arising from hepatocytes or intrahepatic bile duct epithelial cells, which is one of the common malignancies worldwide. Better identification of liver cancer biomarkers has become one of the current challenges. Although hypoxia inducible lipid droplet associated (HILPDA) has been reported to be associated with tumor progression in a variety of human solid cancers, it has rarely been reported in the field of hepatocellular carcinoma; therefore, in this paper, RNA sequencing data from TCGA were used to analyze the expression of HILPDA and differentially expressed genes (DEGs). In addition, functional enrichment analysis of HILPDA-associated DEGs was performed by GO/KEGG, GSEA, immune cell infiltration analysis and protein-protein interaction network. The clinical significance of HILPDA in LIHC was calculated by Kaplan–Meier Cox regression and prognostic nomogram models. R package was used to analyze the combined studies. Thus, HILPDA was highly expressed in various malignancies, including LIHC, compared with normal samples, and high HILPDA expression was associated with poor prognosis ( $P < .05$ ). Cox regression analysis showed high HILPDA to be an independent prognostic factor; age and cytogenetic risk were included in the nomogram prognostic model. A total of 1294 DEGs were identified between the high and low expression groups, of which 1169 had upregulated gene expression and 125 had downregulated gene expression. Overall, high expression of HILPDA is a potential biomarker for poor outcome in LIHC.

**Abbreviations:** AFP = alpha-fetoprotein, DEGs = differentially expressed genes, GO = gene ontology, HILPDA = hypoxia inducible lipid droplet associated, KEGG = kyoto encyclopedia of genes and genomes, LIHC = hepatocellular carcinoma, OS = overall survival, PPI = protein-protein interaction, RNA-seq = RNA sequencing, ROC = receiver operating characteristic curve, ssGSEA = single sample gene set enrichment analysis, TCGA = the cancer genome atlas.

**Keywords:** bioinformatics, hepatocellular carcinoma, HILPDA, immunohistochemical, R package

## 1. Introduction

Hepatocellular carcinoma (LIHC) is a kind of malignant tumor that occurs from hepatocytes or intrahepatic bile duct epithelial cells, which is one of the common malignant tumors worldwide. According to the latest statistics published by the International Agency for Research on Cancer, the number of new cases of liver cancer worldwide in 2020 is about 906,000 and the number of deaths is about 830,000, and the incidence and mortality of liver cancer occupy 4.7% and 8.3% of the total cancer incidence and mortality, respectively, ranking 6th and 3rd among all tumors.<sup>[1]</sup> Considering the complexity of liver cancer tumor biology and the limitations of various therapeutic approaches, it has become a consensus in the field of liver cancer treatment to apply multidisciplinary collaboration

and coexistence of multiple therapeutic approaches for comprehensive treatment. In recent years, high-level clinical research evidence is emerging and comprehensive treatment strategies for hepatocellular carcinoma are being developed and refined, during which the early diagnosis of hepatocellular carcinoma has been improved. However, due to its insidious onset and lack of significant symptoms in the early stage, most diagnosed patients have progressed to the middle and advanced stages, which makes it difficult to perform effective surgical treatment, thus resulting in a low 5-year survival rate of hepatocellular carcinoma patients.<sup>[2]</sup> Therefore, an in-depth study of the molecular mechanisms of hepatocarcinogenesis and progression is important for finding new diagnostic and treatment markers to improve the prognosis of liver cancer patients.

This work was supported by grants from the National Key R&D Program of China (no. 2019YFC1712500), National Natural Science Foundation of China (no. 82060163), the Science and Technology Department of Guizhou Province (QKHHBZ[2020]3003).

The authors have no conflicts of interest to disclose.

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

The data for this study were obtained from databases and did not require ethical review.

Supplemental Digital Content is available for this article.

<sup>a</sup> Guizhou University of Traditional Chinese Medicine, Guiyang, China, <sup>b</sup> Guangzhou University of Traditional Chinese Medicine, Guangzhou, China.

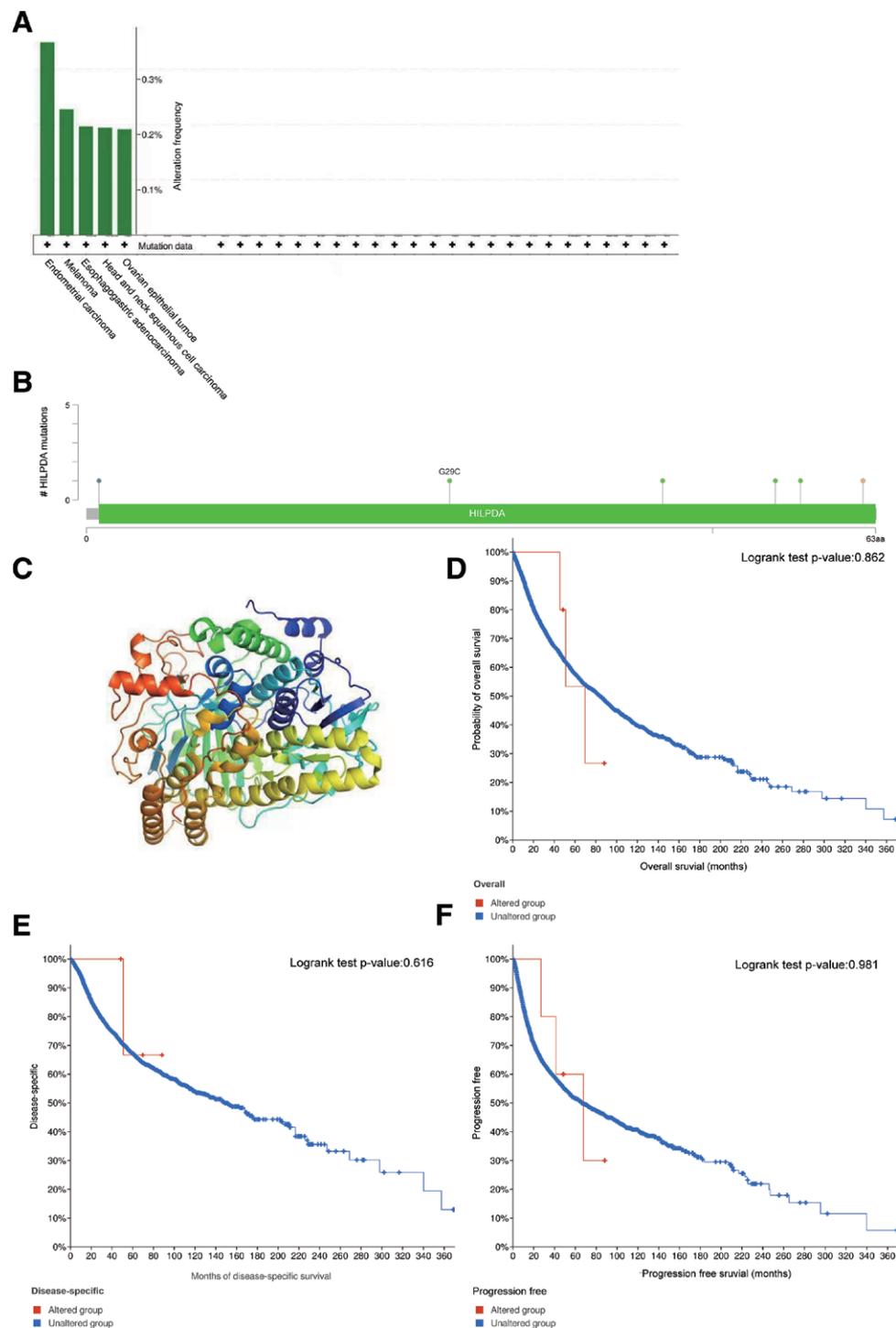
\*Correspondence: Xia Liu, Guizhou University of Traditional Chinese Medicine, Guiyang, Guizhou 550025, China (e-mail: liuxia0851@126.com).

Copyright © 2023 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 X, Zou A, Zhang J, Gao G, Shan W, Li J, Liu X. High expression of HILPDA is an adverse prognostic prognostic factor in hepatocellular carcinoma. *Medicine* 2023;102:9(e33145).

Received: 11 May 2022 / Received in final form: 7 February 2023 / Accepted: 10 February 2023

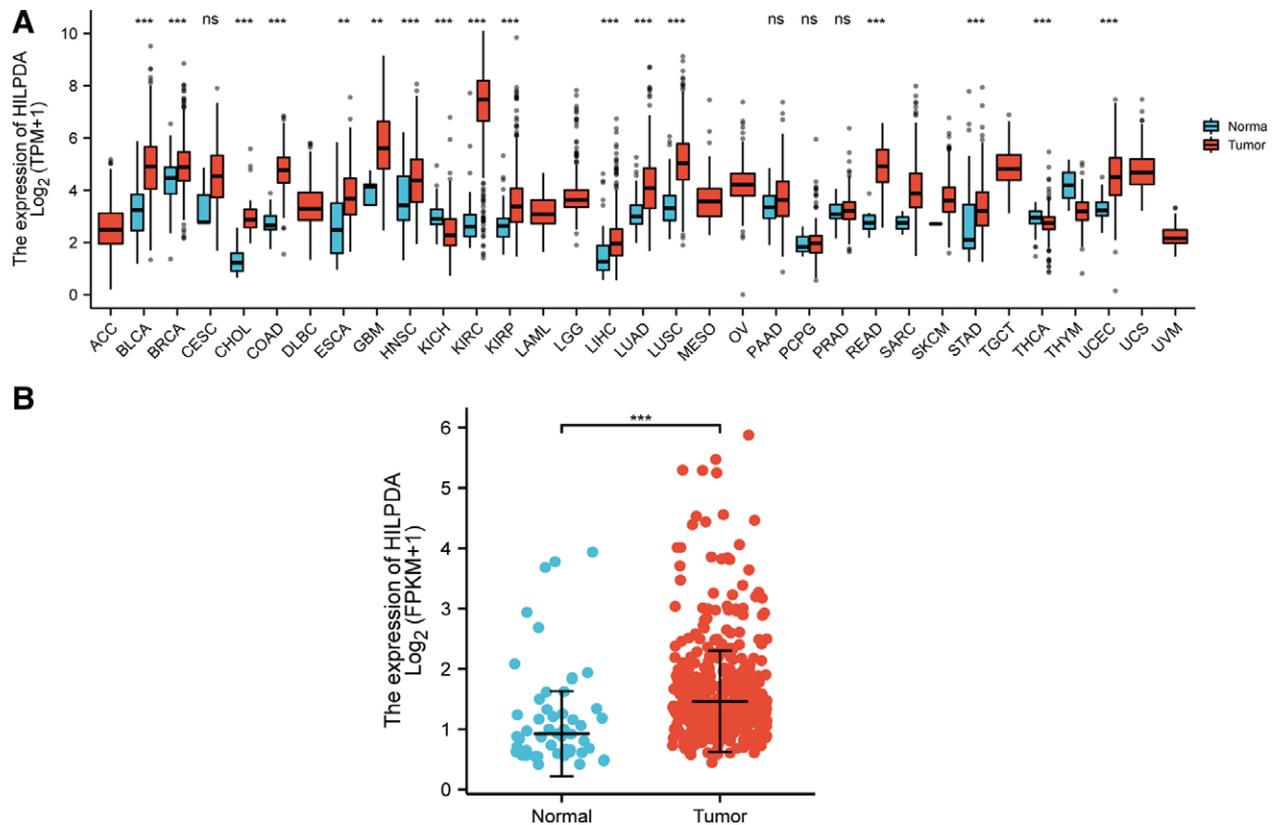
<http://dx.doi.org/10.1097/MD.00000000000033145>



**Figure 1.** Characteristics of HILPDA mutation in TCGA pan-cancer cohort and 3D protein structure of HILPDA. (A) Prevalence of HILPDA mutations in tumors. (B) Subtypes and distribution of HILPDA somatic mutations. X-axis indicates amino acids; Y-axis indicates the number of HILPDA mutations. (C) 3D protein structure of HILPDA. (D) Overall survival (OS) analysis stratified by HILPDA mutation status in the entire TCGA cohort. (E) Disease-specific survival (DSS) analysis stratified by HILPDA mutation status in TCGA. (F) Progression-free survival (PFS) analysis stratified by HILPDA mutation status in TCGA. ns,  $P > .05$ ; \* $P < .05$ ; \*\*\* $P < .001$ . HILPDA = hypoxia inducible lipid droplet associated, TCGA = the cancer genome atlas.

Hypoxia inducible lipid droplet associated (HILPDA) is located on chromosome 7 and consists of two exons and one intron covering a 2.6 kb genomic region.<sup>[3]</sup> It is a hypoxia-inducible lipid droplet-associated protein. HILPDA is located in the intracellular endoplasmic reticulum and around lipid droplets, and can affect lipid storage in hepatocytes, macrophages and cancer cells.<sup>[4]</sup> It has been reported that HILPDA can affect the prognosis of cancer patients by participating in the hypoxic state

and lipid metabolism in cancer patients.<sup>[5]</sup> Therefore, the liver, as an important metabolic organ in the body, may be highly susceptible to HILPDA. Although the role of HILPDA in immune cells is unclear, it has been shown to be regulated by progesterone in a hypoxia-dependent manner in cancer<sup>[6]</sup> and to play a carcinogenic role in a variety of tumor types. For example, the imbalance of HILPDA is related to the biological effects of proliferation, movement, apoptosis and iron metabolism of hepatoma cells in



**Figure 2.** Pan-cancer analysis of HILPDA and its expression in LIHC. (A) Expression levels of HILPDA in paired normal and pan-cancer samples. (B) Expression levels of HILPDA in paired normal and LIHC samples. ns:  $P \geq .05$ ; \* $P < .05$ ; \*\* $P < .01$ ; \*\*\* $P < .001$ . HILPDA = hypoxia inducible lipid droplet associated, LIHC = hepatocellular carcinoma.

vitro.<sup>[7]</sup> HILPDA is overexpressed in colorectal cancer, and leads to further cancer development through hypoxia-dependent and independent pathways.<sup>[8]</sup> High HILPDA expression predicts poor survival in patients with renal cancer, which may become a potential target for molecular therapy.<sup>[9]</sup> HILPDA is an important autocrine growth factor in Wnt signaling,<sup>[9]</sup> and it is also a predictive marker for the survival of many cancer patients.<sup>[10]</sup>

In this study, we aimed to determine the relationship between HILPDA expression levels and LIHC prognosis through the following four steps: First, the cBioPortal database was used to examine the prevalence of HILPDA somatic mutations in the Cancer Genome Atlas (TCGA) pan-cancer cohort, and to analyze the subtype and distribution of HILPDA in somatic mutations, the variant sites on 3D protein structures, the tumor mutational load in HILPDA non-mutant cancers and different subtypes of HILPDA mutant cancers, and the survival prognosis. Subsequently, RNA sequencing (RNA-seq) data from LIHC samples from TCGA was obtained and genotype-tissue expression were used to analyze the expression of the core gene HILPDA. In addition, functional enrichment analysis of HILPDA was performed by gene ontology (GO), kyoto encyclopedia of genes and genomes (KEGG), GSEA, immune cell infiltration analysis, and protein-protein interaction (PPI) network. Finally, the clinical significance of HILPDA in LIHC was analyzed by Kaplan–Meier and Cox regression and columnar line graph prognostic models.

## 2. Materials and Methods

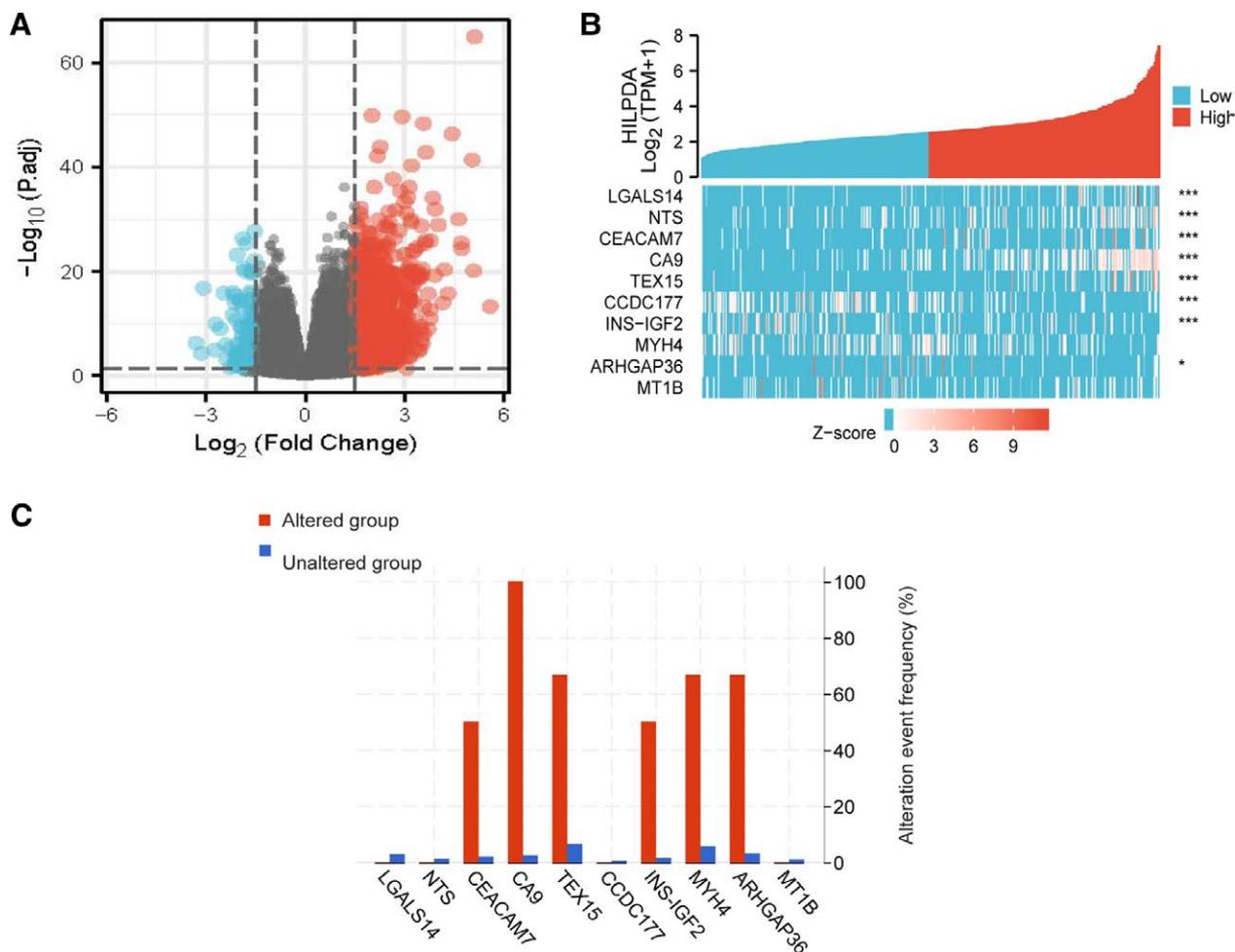
### 2.1. TCGA data

According to the official website of The Cancer Genome Atlas (TCGA), the database includes sequencing and clinicopathological data from patients with more than 30 types of tumors. All

data for prevalence analysis of HILPDA mutations and CNA, subtype analysis, variant sites on 3D protein structures, tumor mutational load and survival analysis in HILPDA non-mutant cancers and different subtypes of HILPDA mutant cancers are available from the cBioportal for Cancer Genomics database (<https://www.cbioportal.org>) for query and download (cbioportal.org).<sup>[11]</sup> The cBioPortal database integrates data from TCGA, CCLE, and several large independent oncology research projects, storing a wide range of tumor histology data that can be queried, analyzed and visualized for relevant results.

### 2.2. RNA sequencing data and bioinformatics analysis

In recent years, the rapid development of sequencing technology has greatly facilitated researchers' understanding of the transcriptional environment and contributed to a deeper and more detailed study of gene expression. RNA sequencing technology has advantages over previous analytical methods, as it can be performed at the single-cell and whole-genome levels, and can detect thousands of samples simultaneously with high throughput and high resolution. Along with the development of sequencing technology, multiple bioinformatics analysis methods have been developed, including quality control and read quantification. UCSC Xena is a cancer genomics data analysis platform that supports the visualization and analysis of multiple histological data from cancer samples, with built-in public datasets, such as data from TCGA, ICGC and other large cancer research projects. It also supports the analysis of your own data, and ensures the security of your data, so that you don't have to worry about your data being stolen by other users after uploading. Uniformly processed pan-cancer RNA-seq data from TCGA were downloaded from UCSC XENA (<https://xenabrowser.net/datapages/>).<sup>[12–15]</sup> Level 3 HTSeq-FPKM and HTSeq-Count data for LIHC samples



**Figure 3.** Differentially expressed genes in LIHC samples at low and high expression of HILPDA. (A) The volcano map of differentially expressed genes included 1169 up-regulated genes and 125 down-regulated genes. Standardized expression levels are shown in descending order from green to red. (B) Heat map of ten differentially expressed RNAs including five up-regulated genes and five down-regulated genes. X-axis represents samples, and Y-axis represents differentially expressed RNAs. The green bars and the red bars represent down- and up-regulated genes, respectively. (C) Ten kinds of genes are represented in HILPDA mutant and non-mutant cancers. HILPDA = hypoxia inducible lipid droplet associated, LIHC = hepatocellular carcinoma.

were obtained from the TCGA website (<https://portal.gdc.cancer.gov/repository>) for further analysis. This study fully complies with the guidelines published by TCGA.

### 2.3. Differentially expressed gene (DEG) analysis

The DESeq2<sup>7</sup> package was used to compare low expression and high expression of HILPDA (cutoff value of 50%) in LIHC samples (HTseq-Count) to identify DEGs.<sup>116</sup> The first 10 DEGs were performed by heat map.

### 2.4. Functional enrichment analysis

The KEGG pathway enrichment analysis is described in terms of biochemical pathways and regulatory pathways of genes. In this paper, functional and pathway enrichment analysis of DEGs was performed using the ClusterProfiler package of the R package, and genes with  $DEG > 1.5$  and  $Padj < .05$  at the  $llogFC$  threshold were screened for functional enrichment analysis. GO functional analysis includes cellular component, molecular function, and biological process.<sup>177</sup>

### 2.5. Gene set enrichment analysis (GSEA)

The R package ClusteProfiler (3.14.3) was used for GSEA to elucidate the functional and pathway differences between the

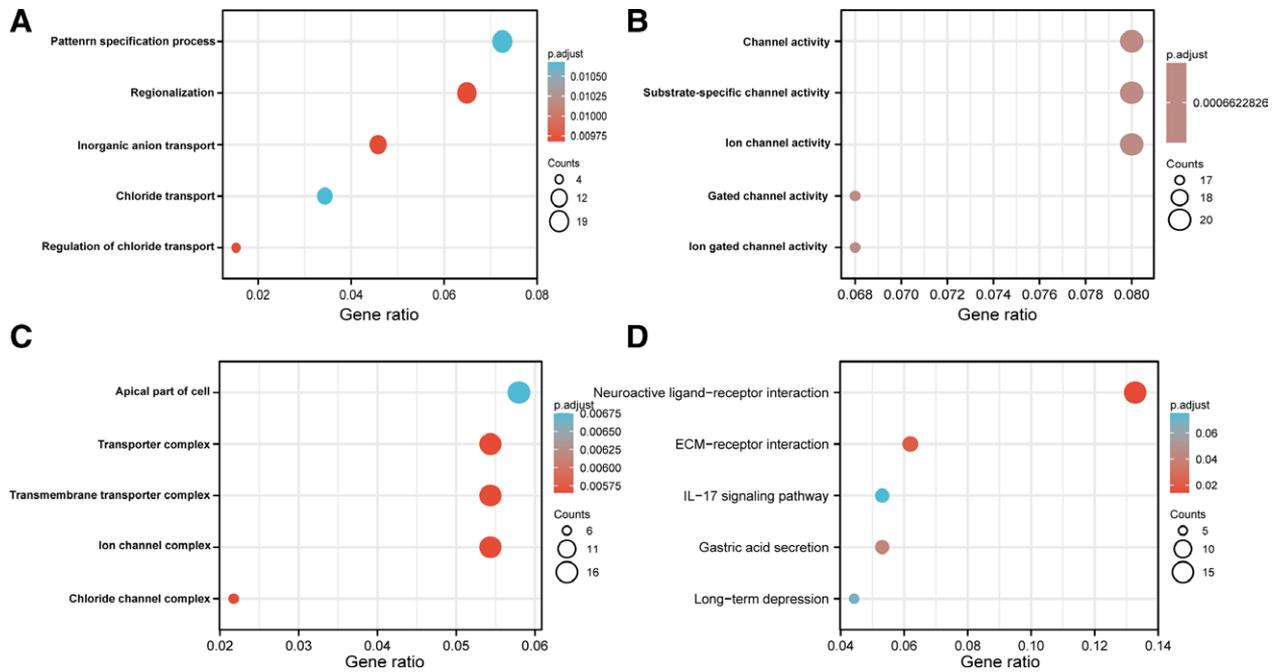
high and low expression groups of HILPDA.<sup>7</sup> Gene sets were permuted 1000 times after each analysis.  $P < .05$  and FDR  $q$ -values  $< 0.25$  after permutation were considered statistically significant.

### 2.6. Immuno-infiltration analysis by single sample gene set enrichment analysis (ssGSEA)

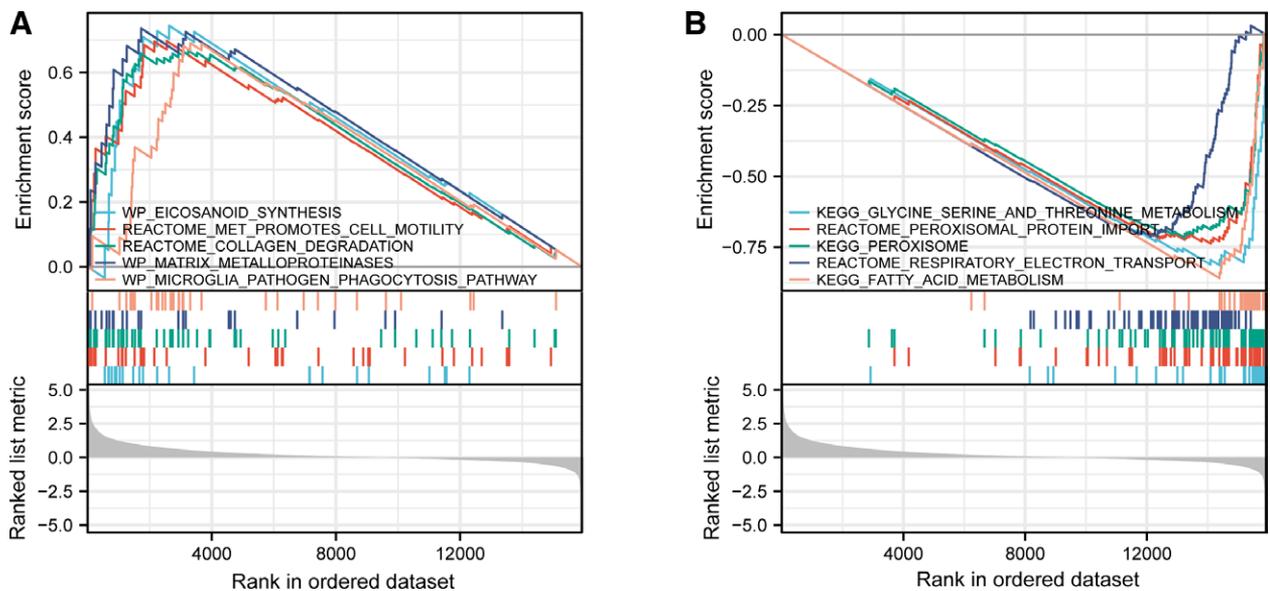
As an extension of the GSEA method, ssGSEA defines an enrichment score such that the score represents the absolute enrichment of gene sets in each sample within a given dataset, and ssGSEA is implemented by the R package GSVA. Immuno-infiltration analysis of HILPDA was performed by ssGSEA using the GSVA package in R (3.6.3). A total of 24 infiltrating immune cells were obtained as described previously.<sup>181</sup> Spearman correlation was used to analyze the correlation between HILPDA and 24 immune cell enrichment scores. Wilcoxon rank sum test was used to analyze the enrichment scores of high and low HILPDA expression groups.

### 2.7. PPI network

The gene sets obtained in this paper were enriched in cancer-related pathways, and most of the genes were related to the occurrence and development of cancer. To further obtain the



**Figure 4.** GO/KEGG enrichment analysis of DEG between high and low HILPDA expression in TCGA-LIHC patients. (A) Enriched GO terms in the “Biological Processes” (BP) category; (B) Enriched GO terms in the “Molecular Functions” (MF) category. (C) Enriched GO terms in the “Cellular Component” (CC) category; (D) KEGG pathway annotation. X-axis represents the proportion of DEG, and Y-axis represents different categories. Different colors represent different properties and different sizes represent the number of DEGs. DEGs = differentially expressed genes, GO = gene ontology, HILPDA = hypoxia inducible lipid droplet associated, KEGG = kyoto encyclopedia of genes and genomes, LIHC = hepatocellular carcinoma, TCGA = the cancer genome atlas.



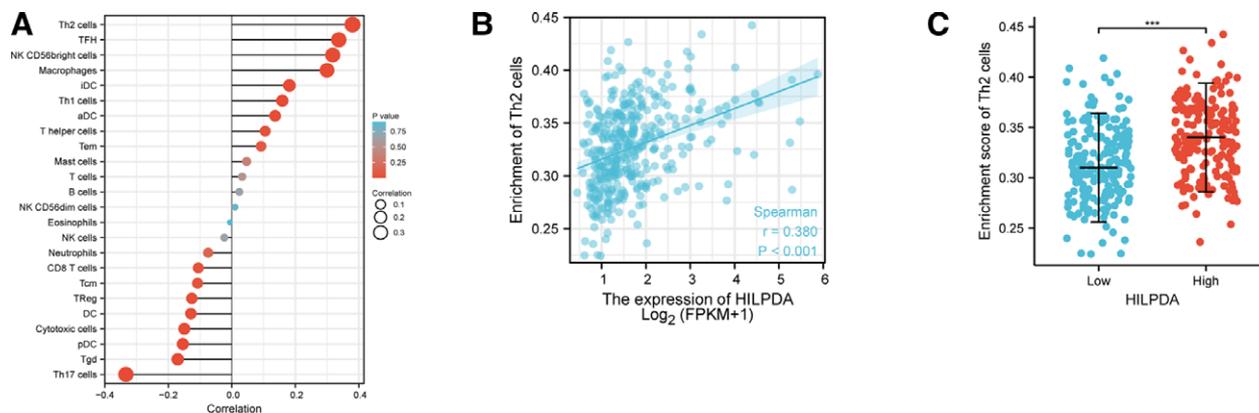
**Figure 5.** Single gene set enrichment analysis (GSEA) enrichment map. (A) Enrichment plots of positively correlated five-GSEA. (B) Enrichment plot of negatively correlated five-GSEA. KEGG = kyoto encyclopedia of genes and genomes.

Hub genes; all the genes were entered into the STRING database to obtain the PPI network of gene interactions. Based on the protein-protein interaction data obtained from the STRING database, the PPI network for the DEGs in this paper was also predicted using the Search Tool for Retrieval of Interacting Genes (STRING) database.<sup>[19]</sup> A 0.4 interaction score threshold was set as the cutoff criterion. the PPI network was mapped using Cytoscape (version 3.7.1)<sup>[20]</sup> and the most important modules in the PPI network were identified using MCODE (version 1.6.1).<sup>[21]</sup> The selection criteria were as follows: MCODE score > 5, degree cutoff = 2, node score cutoff = 0.2, maximum depth

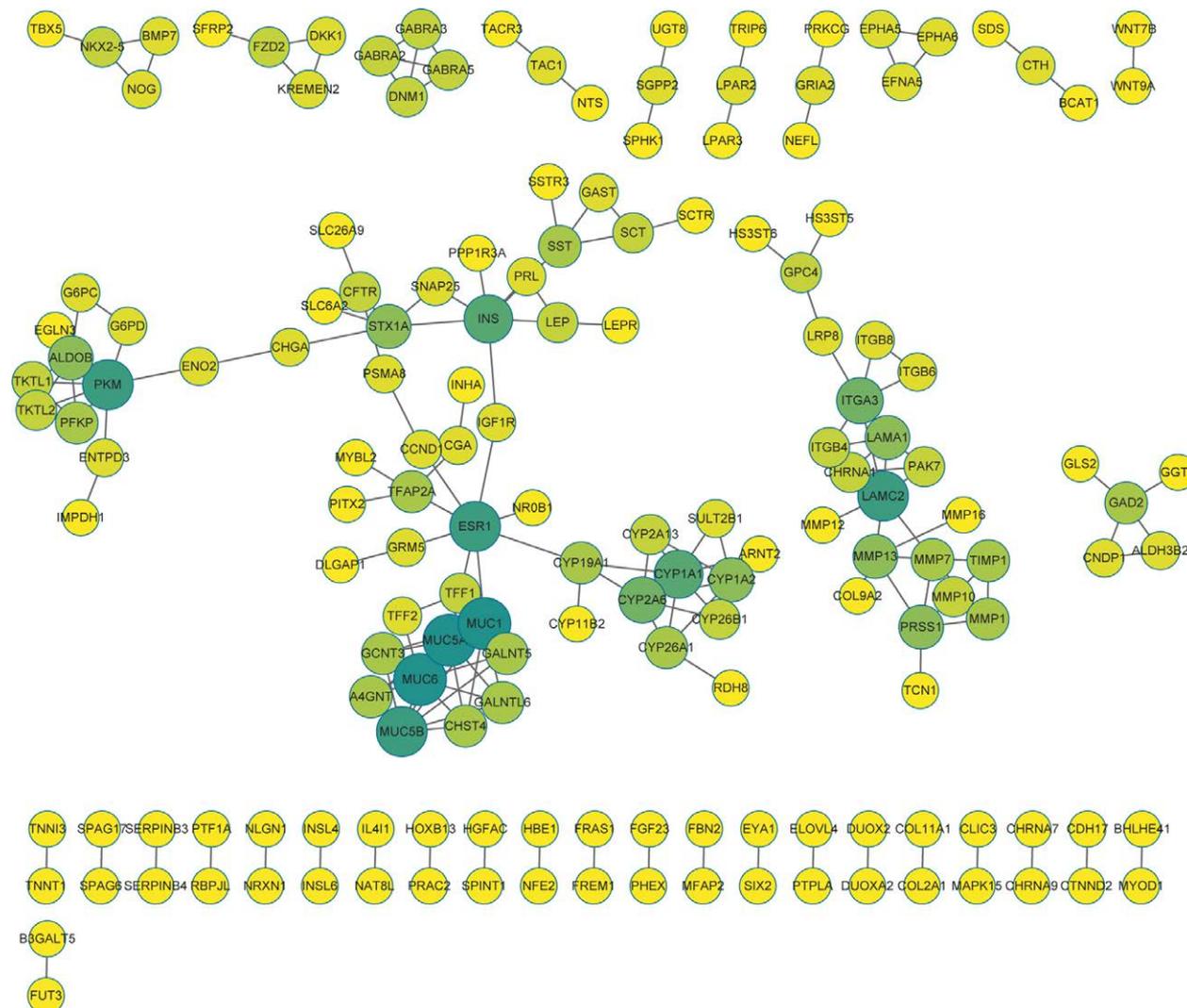
= 100, k-score = 2. Metascape (<https://metascape.org/gp/index.htm>) was used to perform pathway and process enrichment analysis.

### 2.8. Prognostic model generation and prediction

To personalize the prediction of overall survival (OS) and event-free survival in patients with LIHC, column line plots were generated using the RMS R package (version 5.1-3), which includes salient clinical features and calibration plots. Calibration curves were evaluated graphically by mapping



**Figure 6.** Correlation analysis of HILPDA expression with immune infiltration in the LIHC microenvironment. (A) Forest plots showed a positive correlation between HILPDA and 13 immune cells and a negative correlation between HILPDA and 11 immune cell subpopulations. The size of the dots indicates the absolute value of Spearman  $r$ . (B) Correlation between the relative enrichment score of Th2 (bright) cells and the expression level of HILPDA (TPM). (C) Infiltration of Th2 (bright) cells between low and high expression of HILPDA. HILPDA = hypoxia inducible lipid droplet associated.



**Figure 7.** PPI network of HILPDA-related DEGs. DEGs = differentially expressed genes, HILPDA = hypoxia inducible lipid droplet associated, PPI = protein-protein interaction.

the probabilities predicted by the column line plots to the observed ratios, with the 45° line representing the best predicted value. The consistency index was used to determine the

discrimination of the column line plots and 1000 resamples were calculated using the bootstrap method. Patients with hepatocellular carcinoma were divided into high-risk and

**Table 1****Association between HILPDA expression and clinicopathological features in LIHC samples from the TCGA database.**

Characteristic	Low expression of HILPDA	High expression of HILPDA	P
n	187	187	
Age, n (%)			
≤60	83 (22.3%)	94 (25.2%)	.323
>60	103 (27.6%)	93 (24.9%)	
Gender, n (%)			
Female	57 (15.2%)	64 (17.1%)	.507
Male	130 (34.8%)	123 (32.9%)	
Pathologic stage, n (%)			
Stage I	109 (31.1%)	64 (18.3%)	<.001
Stage II	35 (10%)	52 (14.9%)	
Stage III	30 (8.6%)	55 (15.7%)	
Stage IV	1 (0.3%)	4 (1.1%)	
T stage, n (%)			
T1	116 (31.3%)	67 (18.1%)	<.001
T2	36 (9.7%)	59 (15.9%)	
T3	29 (7.8%)	51 (13.7%)	
T4	3 (0.8%)	10 (2.7%)	
N stage, n (%)			
N0	123 (47.7%)	131 (50.8%)	.124
N1	0 (0%)	4 (1.6%)	
AFP (ng/mL), n (%)			
≤400	123 (43.9%)	92 (32.9%)	.227
>400	31 (11.1%)	34 (12.1%)	
Fibrosis ishak score, n (%)			
0	44 (20.5%)	31 (14.4%)	.142
1/2	13 (6%)	18 (8.4%)	
3/4	18 (8.4%)	10 (4.7%)	
5/6	53 (24.7%)	28 (13%)	
OS event, n (%)			
Alive	144 (38.5%)	100 (26.7%)	<.001
Dead	43 (11.5%)	87 (23.3%)	
DSS event, n (%)			
Alive	160 (43.7%)	127 (34.7%)	<.001
Dead	25 (6.8%)	54 (14.8%)	
PFI event, n (%)			
Alive	103 (27.5%)	88 (23.5%)	.148
Dead	84 (22.5%)	99 (26.5%)	
Child-Pugh grade, n (%)			
A	128 (53.1%)	91 (37.8%)	.105
B	8 (3.3%)	13 (5.4%)	
C	1 (0.4%)	0 (0%)	
Vascular invasion, n (%)			
No	120 (37.7%)	88 (27.7%)	.023
Yes	48 (15.1%)	62 (19.5%)	
Age, median (IQR)	62 (53, 69)	60 (51, 69)	.207

AFP = alpha-fetoprotein, DSS = disease-specific survival, HILPDA = hypoxia inducible lipid droplet associated, LIHC = hepatocellular carcinoma, OS = overall survival, TCGA = the cancer genome atlas.

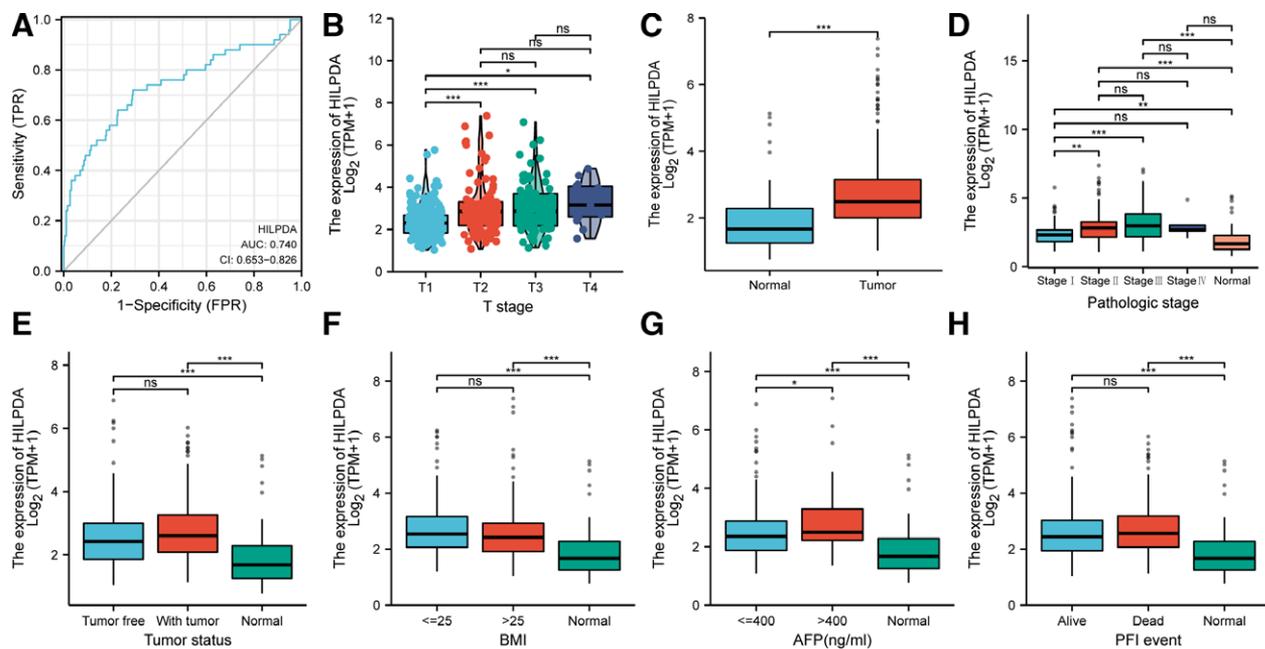
low-risk groups according to the median value of risk scores. Kaplan–Mteier survival curves were applied to compare the survival rates (overall survival, OS) between them. A time-dependent receiver operating characteristic curve (ROC) analysis was performed using the “survival ROC” package to assess the credibility of the model. Finally, we used univariate and multifactorial cox regression analyses to combine risk scores with clinical information to assess the relationship with patient prognosis and to identify independent prognostic factors. All statistical tests were performed using a two-tailed test with 0.05 as the level of statistical significance.

### 2.9. Immunohistochemical validation

The immunohistochemical images were found by HPA database, accessing HPA database <https://www.proteinatlas.org/>, searching Aregap36, Ca9, Ceacam7, Hilpda, Ins-Igf2, Myh4, Nts, Tex15 to get eight genes in normal tissues and LIHC patients Differences.

### 2.10. Statistical analysis

All statistical analyses and graphs were analyzed and displayed by R (3.6.2).<sup>[22]</sup> Cox regression analysis and Kaplan–Meier method were used to assess prognostic factors. Survival analysis was performed using the Kaplan–Meier method, and HILPDA expression in unpaired samples was analyzed using the Wilcoxon rank sum test, and paired samples were analyzed using the Wilcoxon signed rank test. Multivariate Cox analysis was used to compare the effect of HILPDA expression on survival as well as other clinical characteristics. Median HILPDA expression was considered as the cutoff value. A  $P < .05$  was considered statistically significant in all tests. In addition, ROC analysis was performed in the PROC package to assess the validity of transcript expression of HILPDA in differentiating LIHC from healthy samples. The calculated area under the curve values range from 0.5 to 1.0, indicating a 50% to 100% discrimination capability.



**Figure 8.** Association between HILPDA expression and clinical features. (A) The diagnostic efficacy of HILPDA in hepatocellular carcinoma was analyzed by ROC. (B–H) The association between HILPDA expression and T stage, status, pathologic stage, Tumor status, BMI, AFP, PFI was analyzed using the Wilcoxon Rank SUM test. AFP = alpha-fetoprotein, BMI = body mass index, HILPDA = hypoxia inducible lipid droplet associated, PFI = platinum-free interval, ROC = receiver operating characteristic curve, TPM = transcripts per kilobase of exon model per Million mapped reads, TPR = true positive rate.

**Table 2**

**Logistic analysis to explore the relationship between clinicopathological factors of LIHC and HILPDA expression.**

Characteristics	Total (N)	Odds ratio (OR)	P value
T stage (T3 & T4 vs T1 & T2)	371	2.300 (1.420–3.782)	<.001
N stage (N1 vs N0)	258	70927270.804 (0.000–NA)	.995
M stage (M1 vs M0)	272	2.743 (0.346–55.834)	.385
Pathologic stage (stage III & stage IV vs stage I & stage II)	350	2.363 (1.444–3.926)	<.001
Tumor status (with tumor vs tumor free)	355	1.556 (1.021–2.379)	.040
Age (>60 vs ≤60)	373	0.797 (0.530–1.197)	.275
BMI (>25 vs ≤25)	337	0.793 (0.516–1.217)	.289
AFP (ng/ml) (>400 vs ≤400)	280	1.466 (0.840–2.567)	.178
Vascular invasion (yes vs no)	318	1.761 (1.107–2.818)	.017
Prothrombin time (>4 vs ≤4)	297	1.156 (0.702–1.902)	.568

AFP = alpha-fetoprotein, HILPDA = hypoxia inducible lipid droplet associated, LIHC = hepatocellular carcinoma.

### 3. Results

#### 3.1. Characteristics of HILPDA mutations in the TCGA pan-cancer cohort

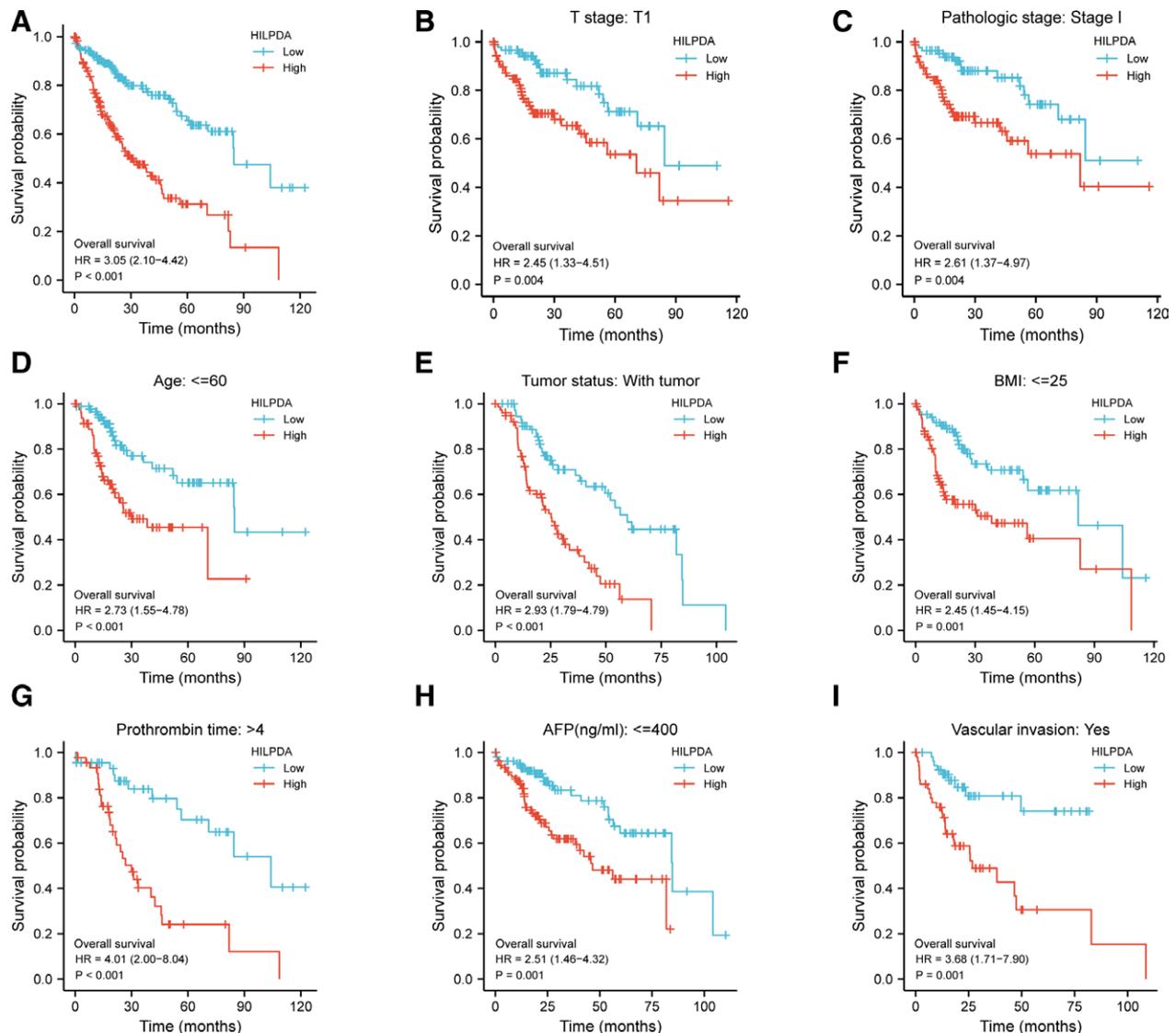
Out of all 2431 patients, 6 (<1%) carried HILPDA mutations (Fig. 1A). HILPDA mutations occurred in a small subset of most tumor types, and the mutation frequency differed significantly across tumors ( $P < .001$ ). A total of 6 HILPDA mutations were identified, 4 (66.6%) were missense mutations, 1 (16.6%) was a truncating mutation and 1 (16.6%) was a splice mutation (Fig. 1B). These mutations occurred in a scattered manner throughout the sequence (Fig. 1B) and the 3D protein structure (Fig. 1C). To investigate whether these different characteristics of HILPDA mutations translate into cancer prognosis, we compared OS ( $P = .862$ , Fig. 1D), disease-specific survival ( $P = .616$ , Fig. 1E), and PFS ( $P = .981$ , Fig. 1F) between HILPDA-mutated cancer patients and HILPDA-nonmutated cancer patients. prognosis and survival of cancer patients in the TCGA cohort were independent of HILPDA mutation status.

#### 3.2. HILPDA expression in pan-cancer and LIHC

RNA-seq data from UCSC XENA (<https://xenabrowser.net/datapages/>) were downloaded in TCGA format and processed uniformly through a cumbersome process. By comparing the expression of HILPDA in normal samples in the TCGA database with the corresponding tumor samples in the TCGA database, it was found that HILPDA was significantly highly expressed in 17 cancers. (Fig. 2A), In particular, the expression of genes in hepatocellular carcinoma (LIHC) showed extremely significant differences (Fig. 2B).

#### 3.3. Differentially expressed genes in LIHC samples at low and high expression of HILPDA

Median mRNA expression differences in gene expression profiles between the high and low expression groups were analyzed. A total of 1294 DEGs from the gene expression RNA-seq-HTSeq-Counts, including 1169 up-regulated and 125 down-regulated, were identified as statistically significant



**Figure 9.** High expression of HILPDA is associated with poorer OS in LIHC patients. (A) Kaplan–Meier curves for all patients with LIHC. (B) Kaplan–Meier curves of LIHC patients with T stage I. (C) Kaplan–Meier curves of LIHC patients with Pathologic stage I. (D) Kaplan–Meier curves for LIHC patients aged  $\leq 60$  years. (E) Kaplan–Meier curves of LIHC patients with tumor. (F) Kaplan–Meier curves for LIHC patients with BMI  $\leq 25$  kg/m<sup>2</sup>. (G) Kaplan–Meier curves for LIHC patients with Prothrombin time  $> 4$  seconds. (H) Kaplan–Meier curves of LIHC patients with AFP  $\leq 400$  ng/mL. (I) Kaplan–Meier curves of LIHC patients with Vascular invasion: Yes. AFP = alpha-fetoprotein, HILPDA = hypoxia inducible lipid droplet associated, LIHC = hepatocellular carcinoma, OS = overall survival.

( $|\log \text{fold change} (\log \text{FC})| > 1.5, P < .05$ ) between the HILPDA high and low expression groups (Fig. 3A). The top five up-regulated DEGs and the top five down-regulated DEGs between the HILPDA high and low expression groups are illustrated by the heat map (Fig. 3B). Ten kinds of genes are represented in HILPDA mutant and non-mutant cancers (Fig. 3C).

### 3.4. Functional enrichment analysis of differentially expressed genes

To better understand the functional significance of 1294 DEG between high and low HILPDA expression in LIHC, GO and KEGG functional enrichment analysis was performed by clusterProfiler package (Table S1, Supplemental Digital Content, <http://links.lww.com/MD/I568>, Fig. 4). Associations with biological processes include pattern specification process, regionalization, inorganic anion transport, chloride transport, regulation of chloride transport; The cellular component includes the apical part of cell transporter complex,

transmembrane transporter complex, ion channel complex, chloride channel complex; Molecular function includes channel activity, substrate-specific channel activity, ion channel activity, gated channel activity, ion gated channel activity. KEGG includes Neuroactive ligand-receptor interaction, ECM-receptor interaction, IL-17 signaling pathway, Gastric acid secretion, Long-term depression (Fig. 4).

### 3.5. Gene set enrichment analysis (GSEA)

GSEA analysis was performed to further understand the biological pathways involved in LIHC with different levels of HILPDA expression. GSEA was performed between low and high HILPDA expression datasets to identify the key signaling pathways involved in LIHC. Significant differences ( $\text{FDR} < 0.25, \text{ADJ } P < .05$ ) were observed in the enrichment of the MSigDB set (C2.all.v7.0.symbols.gmt) for these pathways. NES values were taken in descending order for the top five positive and the top five negative (Fig. 5).

**Table 3****Univariate and multivariate cox regression analysis of factors associated with OS in LIHC.**

Characteristics	Total (N)	HR (95% CI) Univariate analysis	P value Univariate analysis	HR (95% CI) Multivariate analysis	P value Multivariate analysis
T stage	370				
T1 & T2	277	Reference			
T3 & T4	93	2.598 (1.826–3.697)	<.001	2.340 (0.318–17.226)	.404
N stage	258				
N0	254	Reference			
N1	4	2.029 (0.497–8.281)	.324		
M stage	272				
M0	268	Reference			
M1	4	4.077 (1.281–12.973)	.017	1.347 (0.405–4.477)	.627
BMI	336				
≤25	177	Reference			
>25	159	0.798 (0.550–1.158)	.235		
Age	373				
≤60	177	Reference			
>60	196	1.205 (0.850–1.708)	.295		
Pathologic stage	349				
Stage I & Stage II	259	Reference			
Stage III & Stage IV	90	2.504 (1.727–3.631)	<.001	1.112 (0.152–8.155)	.917
Gender	373				
Female	121	Reference			
Male	252	0.793 (0.557–1.130)	.200		
AFP (ng/mL)	279	1.000 (1.000–1.000)	.431		
HILPDA	373				
High	187	Reference			
Low	186	0.333 (0.230–0.482)	<.001	0.334 (0.206–0.540)	<.001

AFP = alpha-fetoprotein, CI = confidence interval, HILPDA = hypoxia inducible lipid droplet associated, LIHC = hepatocellular carcinoma, OS = overall survival.

### 3.6. Immuno-infiltration analysis in LIHC

Spearman correlation analysis showed that the level of HILPDA expression in the LIHC microenvironment correlated with the level of immune cell infiltration quantified by SSGSEA. Specifically, HILPDA was positively correlated with Th2 cells and activated dendritic cells (Fig. 6).

### 3.7. PPI analysis in hepatocellular carcinoma

The network of HILPDA and its potential co-expressed genes in HILPDA-associated DEGs was constructed by STRING with a threshold of 0.4. A total of 1212 DEGs were screened ( $|\log \text{fold change} (\log \text{FC})| > 1.5$ ,  $P < .05$ ). Cytoscape-MCODE shows a PPI network with 164 nodes and 187 edges (Fig. 7). Meanwhile, Metascape-MCODE was used to determine the densely connected PPI network components of HILPDA (Fig. 7).

**3.7.1. Association between HILPDA expression and clinical features and cytogenetic risk.** The main clinical features of LIHC in TCGA are shown in the basic information table. A total of 374 cases (121 women and 253 men) were analyzed in this study. Among them, HILPDA expression was low in 187 (50%) LIHC patients and high in remain. Correlation analysis showed that HILPDA expression was significantly associated with cytogenetic risk and pathological stage, T-stage, overall survival (OS), and disease-specific survival ( $P < .001$ ). High HILPDA expression was significantly increased in both pathological stage and T-stage II, III, and IV, while mortality was significantly higher in overall survival and disease-specific stage with high HILPDA expression than with low HILPDA expression. In addition, HILPDA expression was also strongly associated with other factors, including N stage ( $P = .124$ ), alpha-fetoprotein (AFP) ( $P = .227$ ), fibrosis ishak score ( $P = .142$ ), survival endpoint (PFI) ( $P = .148$ ), liver function grade (Child-Pugh grade) ( $P = .105$ ), vascular invasion ( $P = .023$ ), and fatty liver examination (IQR) ( $P = .207$ ) (Table 1).

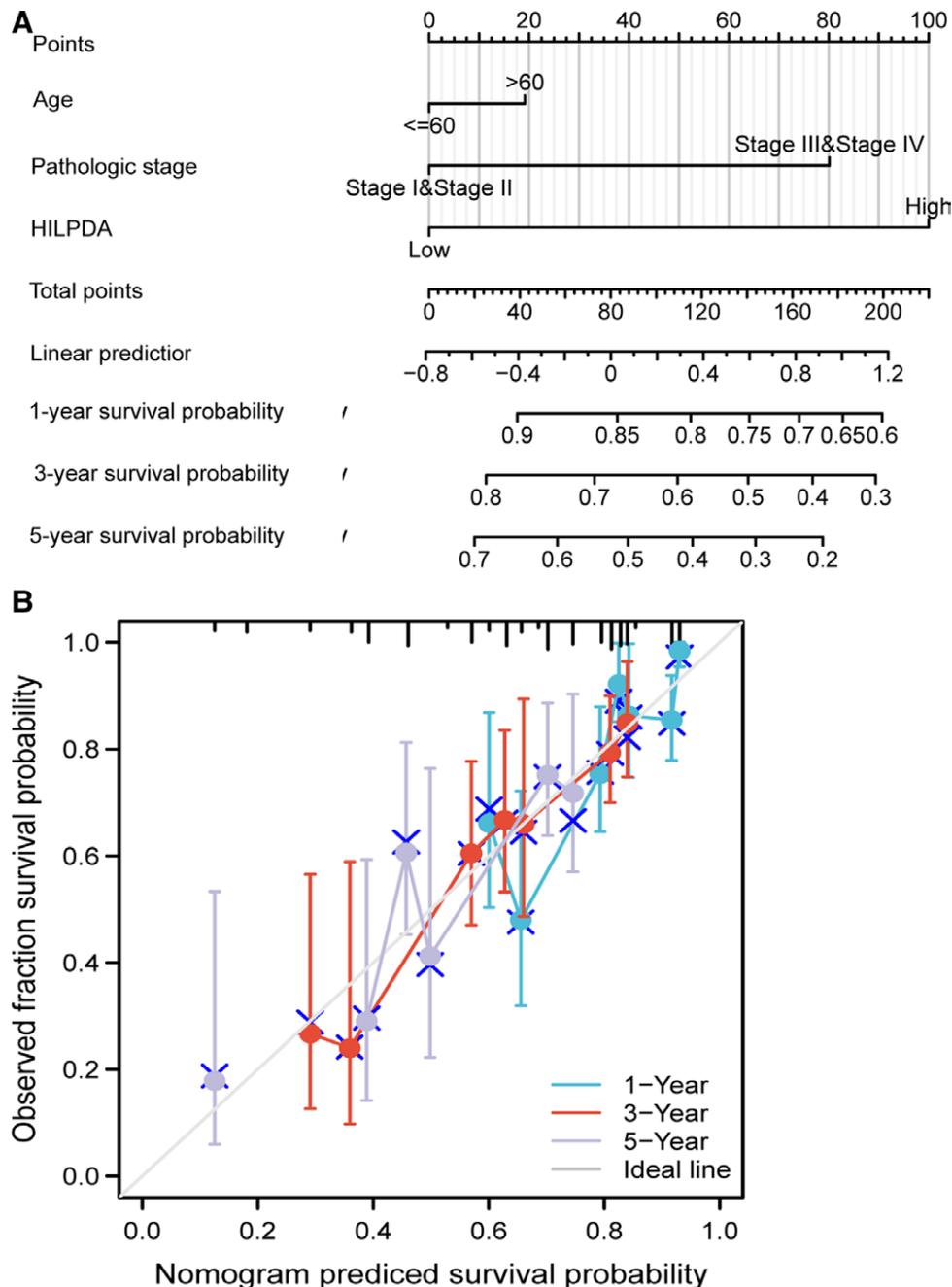
### 3.7.2. Logistic-based analysis of the relationship between clinicopathological factors and HILPDA expression in LIHC.

Logistic analysis was applied to further validate the relationship between LIHC clinicopathological factors and high and low HILPDA dichotomy. As a result, high expression of HILPDA was positively correlated with T-stage (dominance ratio [OR], 2.30;  $P < .001$ ) and pathological stage (dominance ratio [OR], 2.36;  $P < .001$ ), and negatively correlated with AFP (ng/mL) (dominance ratio [OR], 1.46;  $P = .178$ ). In addition, the potential value of HILPDA in differentiating LIHC patients from healthy individuals was examined by ROC curve analysis with an area under the curve of 0.740 (Fig. 8A), indicating the potential of HILPDA as a biomarker. In addition, the Wilcoxon Rank SUM test was used to compare the expression of HILPDA in patients with different clinicopathological characteristics. The results showed that HILPDA was expressed in patients with clinical presentation ( $P < .001$ ), PFI ( $P < .001$ ), AFP ( $P < .001$ ), T stage I and T stage II in T staging ( $P < .001$ ), T stage I and T stage III ( $P < .001$ ), stage I and II in Pathologic ( $P = .005$ ), stage I and III ( $P < .001$ ), Tumor stus ( $P < .001$ ), and BMI ( $P < .001$ ), all of which were statistically significant (Fig. 8B–H; Table 2).

### 3.7.3. High HILPDA affects the prognosis of LIHC in patients with different clinicopathological states.

Kaplan–Meier analysis was applied to analyze the relationship between HILPDA expression and prognosis in patients with LIHC. As seen in Figure 8A, patients with high HILPDA expression had a worse prognosis than those with low HILPDA expression (risk ratio [HR], 3.05 (2.10–4.42);  $P < .001$ ). Kaplan–Meier analysis showed that high HILPDA expression was associated with T stage I ( $P = .004$ ), T Pathologic stage I ( $P = .004$ ), age  $\leq 60$  ( $P < .001$ ), Tumor status of with tumor ( $P < .001$ ), BMI  $\leq 25$  ( $P < .001$ ), Prothrombin time  $> 4$  ( $P < .001$ ), AFP  $\leq 400$  ( $P < .001$ ), Vascular invasion: Yes ( $P < .001$ ; Fig. 9).

Thereafter, univariate Cox proportional risk regression was used to assess factors affecting OS, revealing that HILPDA (high vs. low,  $P = .003$ ) was a predictor of worse OS, as were T stage

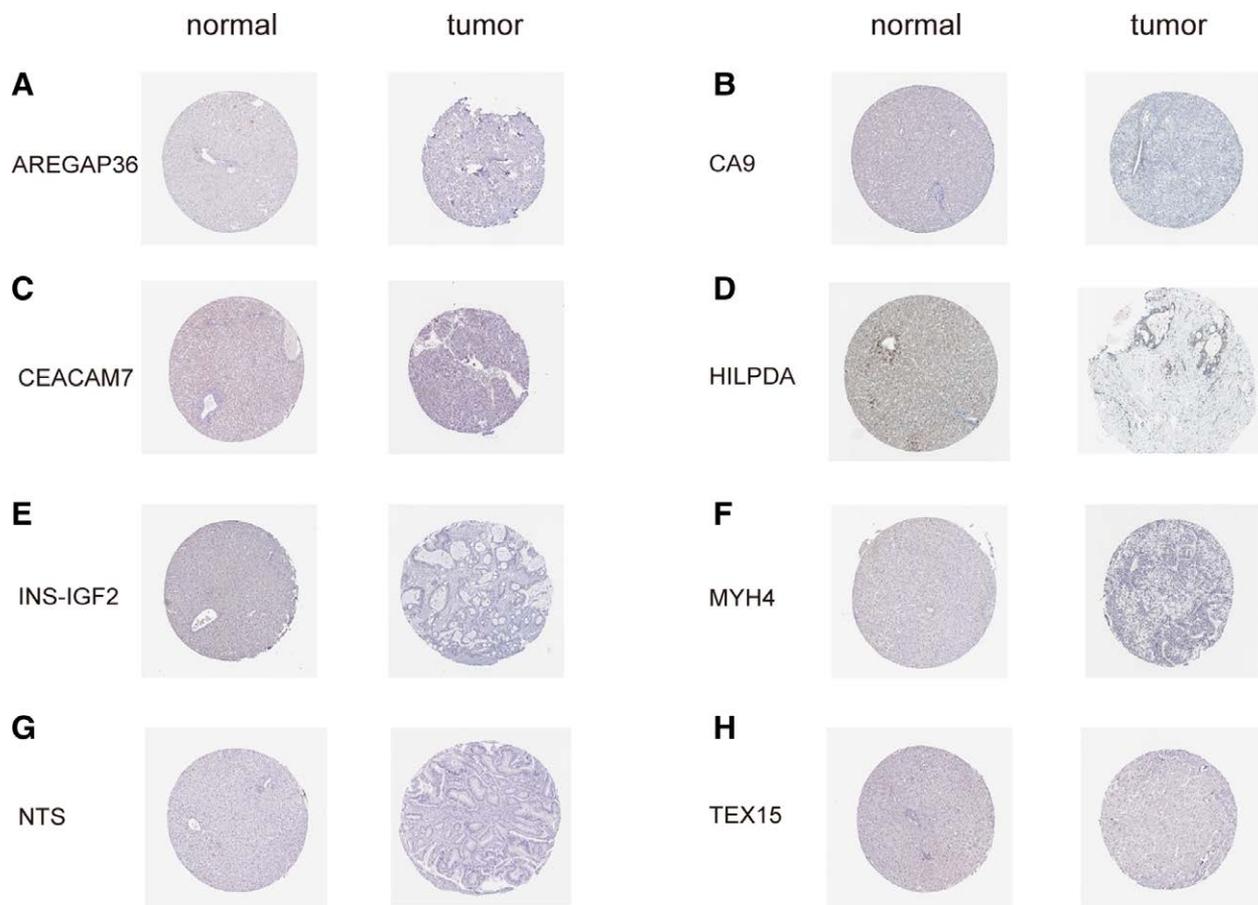


**Figure 10.** Prognostic prediction model of HILPDA in LIHC. (A) Column line graphs for predicting 1-, 3-, and 5-year OS probabilities for LIHC. (B) Columnar line graph calibration plots for predicting 1-, 3-, and 5-year OS probabilities. HILPDA = hypoxia inducible lipid droplet associated, LIHC = hepatocellular carcinoma, OS = overall survival.

III and T IV for T stage ( $P < .001$ ) and stage III and stage IV for Pathologic stage ( $P < .001$ ). T stage III and stage IV, M stage I, Pathologic stage III and stage IV, and HILPDA were then included in multivariate Cox regression, indicating that high HILPDA expression ( $P = .01$ ) was an independent prognostic factor for poorer OS ( $P < .05$ ) (Table 3).

**3.7.4. HILPDA in the LIHC prognostic model.** To better predict the prognosis of patients with LIHC, column line plots were constructed using the RMS R package based on the results of cox regression analysis (Fig. 10A). Three independent prognostic factor variables, age, pathological stage, and HILPDA expression, were included in the model and were selected for inclusion in the prediction model at a statistical

significance level of 0.2. Points were assigned to these variables based on multivariate Cox analysis using a point scale. Lines were drawn upward to determine the number of points for the variables and the sum of points assigned to each variable was rescaled to a range of 0 to 100. The probability of survival for LIHC patients at 1-, 3-, and 5-year was determined by drawing a line from the total point axis straight down to the outcome axis. The 1-year survival probability was determined by drawing a vertical line along the 180 directional end axis on the total point axis, indicating a 1-year survival probability  $< 65\%$ , a 3-year probability  $< 40\%$ , and a 5-year probability even  $< 20\%$ . The predicted results of the OS column line plot calibration curve were consistent with the observations for all patients (Fig. 10B).



**Figure 11.** Immunohistochemistry of AREGAP36, CA9, CEACAM7, HILPDA, INS-IGF2, MYH4, NTS, TEX15 in hepatocellular carcinoma and normal tissues. (A–H) Immunohistochemical maps of eight genes, AREGAP36, CA9, CEACAM7, HILPDA, INS-IGF2, MYH4, NTS, and TEX15 in hepatocellular carcinoma and in normal tissues, respectively.

### 3.8. Immunohistochemical validation of AREGAP36, CA9, CEACAM7, HILPDA, INS-IGF2, MYH4, NTS, TEX15 in hepatocellular carcinoma and normal tissues

Immunohistochemical maps of the top five up-regulated DEGs and the top five down-regulated DEGs between the HILPDA high and low expression groups were selected by prescreening of seven representative genes out of ten genes plus HILPDA for a total of eight genes (Fig. 11).

## 4. Discussion

Liver cancer is one of the common types of cancer and its occurrence is related to several factors, among which chronic hepatitis B virus infection and liver cirrhosis are the most common. Treatments for liver cancer include liver transplantation, liver resection, ablation, chemoembolization, and first-line drugs for targeted treatment of liver cancer including sorafenib, but the treatment effect of liver cancer is still unsatisfactory. Therefore, the discovery of abnormally expressed proteins in hepatocellular carcinoma tissues and the identification of their mechanisms of action will provide new targets of action for the treatment of hepatocellular carcinoma. HILPDA is also associated with multiple disease developments, including several tumor types.<sup>[23]</sup> Previous studies have shown that HILPDA is highly carcinogenic to head and neck cancer,<sup>[24]</sup> mantle cell lymphoma,<sup>[25]</sup> and neuroblastoma.<sup>[26]</sup> However, HILPDA has not been widely studied. Therefore, it is urgent to elucidate the role of HILPDA in tumor progression and treatment. Other studies have shown the relationship between HILPDA expression

and tumor immunosuppressive microenvironment. It was found that HILPDA expression and immunosuppressive genes such as PD-L1, PD-1, TGFB1, and TGFBR1 played a key role in regulating tumor immunosuppressive microenvironment. The high expression of HILPDA suggests that most tumors have immunosuppressive effects, providing potential targets for immunotherapy.<sup>[27]</sup>

Our study showed that HILPDA expression was highly correlated with the prognosis of LIHC patients. Based on the expression of HILPDA in the TCGA database, grouping of LIHC patients revealed significant differences in the expression of genes such as CA9, CEACAM7, INS-IGF2, AREGAP36, MYH4, NTS, and TEX15. This suggests that HILPDA may have an impact on the prognosis of LIHC patients by regulating the above genes. Most of these genes are related to oxidative stress and metabolism. Among them, CA9 is a member of the carbonic anhydrase family, which is usually expressed in cancer cells under hypoxic conditions.<sup>[28]</sup> It controls intracellular pH and protects cancer cells from hypoxia-induced apoptosis.<sup>[29]</sup> In many malignancies, CA9 is highly associated with hypoxia and is regulated by the transcription factor HIF-1 $\alpha$ .<sup>[30]</sup> aberrant methylation of INS-IGF2 is highly susceptible to metabolic disorders associated with breast cancer, pancreatic cancer, diabetes and endocrine-related diseases.<sup>[31]</sup>

In immune cell infiltration analysis, high expression of HILPDA was associated with higher Th2 cells. It was shown that naive T cells met their energy requirements mainly through glycolysis and lipid metabolism<sup>[32,33]</sup> and differentiated into Th2 cells by GATA3 and STAT6. HILPDA-driven differentiation of T cells to Th2 cells may be related to the

compensatory mechanism of human immunity. In contrast, high expression of HILPDA was associated with eicosanoid synthesis, reactive enzymes promoting cell motility reactive enzymes for collagen degradation, matrix metalloproteinases, and lymphoma pathogens, suggesting that HILPDA is not only a potential prognostic biomarker, but also a promising therapeutic target by affecting tumorigenesis-related pathways in LIHC. The deranged expression of HILPDA is highly susceptible to abnormalities in inorganic anion transport such as chloride ions and the activity of transmembrane transporter complexes. KEGG showed that HILPDA interacts with Neuroactive ligand-receptor interaction, ECM-receptor interaction IL-17 signaling pathway, Gastric acid secretion, Long-term depression and other pathways. Notably, the most clinically relevant finding was that high expression of HILPDA was associated with poorer survival. Univariate Cox regression analysis showed that HILPDA expression was associated with T-stage, Stage stage. The development of a column line graph prediction model further confirmed the predictive role of HILPDA expression on prognosis. Therefore, HILPDA may become a new poor prognostic factor for patients with LIHC. In conclusion, the results of this study indicate that high expression of HILPDA in hepatocellular carcinoma tissues is associated with the occurrence and progression of hepatocellular carcinoma and poor prognosis of patients, which identifies the basis for subsequent studies.

## Author contributions

**Conceptualization:** Xiao Wang.

**Data curation:** Xiao Wang, Aoshuang Zou, Jinhe Zhang.

**Formal analysis:** Xiao Wang, Jinhe Zhang.

**Funding acquisition:** Jun Li.

**Methodology:** Xiao Wang, Aoshuang Zou.

**Resources:** Xiao Wang.

**Software:** Xiao Wang, Guochuan Gao.

**Supervision:** Wenting Shan.

**Writing – original draft:** Xiao Wang.

**Writing – review & editing:** Xiao Wang, Xia Liu.

## References

- [1] Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71:209–49.
- [2] Anwanwan D, Singh SK, Singh S, et al. Challenges in liver cancer and possible treatment approaches. *Biochim Biophys Acta Rev Cancer.* 2020;1873:188314.
- [3] Denko N, Schindler C, Koong A, et al. Epigenetic regulation of gene expression in cervical cancer cells by the tumor microenvironment. *Clin Cancer Res.* 2000;6:480–7.
- [4] de la Rosa Rodriguez MA, Kersten S. Regulation of lipid droplet homeostasis by hypoxia inducible lipid droplet associated HILPDA. *Biochim Biophys Acta Mol Cell Biol Lipids.* 2020;1865:158738.
- [5] Cui C, Fu K, Yang L, et al. Hypoxia-inducible gene 2 promotes the immune escape of hepatocellular carcinoma from nature killer cells through the interleukin-10-STAT3 signaling pathway. *J Exp Clin Cancer Res.* 2019;38:229.
- [6] Bray JD, Jelinsky S, Gharje R, et al. Quantitative analysis of gene regulation by seven clinically relevant progestins suggests a highly similar mechanism of action through progesterone receptors in T47D breast cancer cells. *J Steroid Biochem Mol Biol.* 2005;97:328–41.
- [7] Xu Y, Luo X, He W, et al. Long non-coding RNA PVT1/miR-150/HIG2 axis regulates the proliferation, invasion and the balance of iron metabolism of hepatocellular carcinoma. *Cell Physiol Biochem.* 2018;49:1403–19.
- [8] Kim SH, Wang D, Park YY, et al. HIG2 promotes colorectal cancer progression via hypoxia-dependent and independent pathways. *Cancer Lett.* 2013;341:159–65.
- [9] Togashi A, Katagiri T, Ashida S, et al. Hypoxia-inducible protein 2 (HIG2), a novel diagnostic marker for renal cell carcinoma and potential target for molecular therapy. *Cancer Res.* 2005;65:4817–26.
- [10] Zhao H, Ljungberg B, Grankvist K, et al. Gene expression profiling predicts survival in conventional renal cell carcinoma. *PLoS Med.* 2006;3:e13.
- [11] Gao J, Aksoy BA, Dogrusoz U, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal.* 2013;6:p11.
- [12] Vivian J, Rao AA, Nothaft FA, et al. Toil enables reproducible, open source, big biomedical data analyses. *Nat Biotechnol.* 2017;35:314–6.
- [13] Goldman MJ, Craft B, Hastie M, et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat Biotechnol.* 2020;38:675–8.
- [14] Li K, Luo H, Luo H, et al. Clinical and prognostic pan-cancer analysis of m6A RNA methylation regulators in four types of endocrine system tumors. *Aging (Albany NY).* 2020;12:23931–44.
- [15] Wang JD, Zhou HS, Tu XX, et al. Prediction of competing endogenous RNA coexpression network as prognostic markers in AML. *Aging (Albany NY).* 2019;11:3333–47.
- [16] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014;15:550.
- [17] Yu G, Wang LG, Han Y, et al. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS.* 2012;16:284–7.
- [18] Bindea G, Mlecnik B, Tosolini M, et al. Spatiotemporal dynamics of intratumoral immune cells reveal the immune landscape in human cancer. *Immunity.* 2013;39:782–95.
- [19] Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* 2019;47:D607–13.
- [20] Demchak B, Hull T, Reich M, et al. Cytoscape: the network visualization tool for GenomeSpace workflows. *F1000Res.* 2014;3:151.
- [21] 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.
- [22] Isidro-Sánchez J, Akdemir D, Montilla-Bascón G. Genome-wide association analysis using R. *Methods Mol Biol.* 2017;1536:189–207.
- [23] VandeKopple MJ, Wu J, Auer EN, et al. HILPDA regulates lipid metabolism, lipid droplet abundance, and response to microenvironmental stress in solid tumors. *Mol Cancer Res.* 2019;17:2089–101.
- [24] van der Mijn JC, Fu L, Khani F, et al. Combined metabolomics and genome-wide transcriptomics analyses show multiple HIF1 $\alpha$ -induced changes in lipid metabolism in early stage clear cell renal cell carcinoma. *Transl Oncol.* 2020;13:177–85.
- [25] Kuci V, Nordström L, Conrotto P, et al. SOX11 and HIG-2 are cross-regulated and affect growth in mantle cell lymphoma. *Leuk Lymphoma.* 2016;57:1883–92.
- [26] Applebaum MA, Jha AR, Kao C, et al. Integrative genomics reveals hypoxia inducible genes that are associated with a poor prognosis in neuroblastoma patients. *Oncotarget.* 2016;7:76816–26.
- [27] Liu C, Zhou X, Zeng H, et al. HILPDA is a prognostic biomarker and correlates with macrophage infiltration in pan-cancer. *Front Oncol.* 2021;11:597860.
- [28] Fiaschi T, Giannoni E, Taddei ML, et al. Carbonic anhydrase IX from cancer-associated fibroblasts drives epithelial-mesenchymal transition in prostate carcinoma cells. *Cell Cycle.* 2013;12:1791–801.
- [29] Li G, Feng G, Achour Y, et al. MN/CA9 as a novel molecular marker for the detection of cancer. *Expert Opin Med Diagn.* 2007;1:91–7.
- [30] Koukourakis MI, Giatromanolaki A, Danielidis V, et al. Hypoxia inducible factor (HIF1 $\alpha$  and HIF2 $\alpha$ ) and carbonic anhydrase 9 (CA9) expression and response of head-neck cancer to hypofractionated and accelerated radiotherapy. *Int J Radiat Biol.* 2008;84:47–52.
- [31] Suzuki H, Li Y, Dong X, Hassan MM, et al. Effect of insulin-like growth factor gene polymorphisms alone or in interaction with diabetes on the risk of pancreatic cancer. *Cancer Epidemiol Biomarkers Prev.* 2008;17:3467–73.
- [32] Menk AV, Scharping NE, Moreci RS, et al. Early TCR signaling induces rapid aerobic glycolysis enabling distinct acute T cell effector functions. *Cell Rep.* 2018;22:1509–21.
- [33] Lochner M, Berod L, Sparwasser T. Fatty acid metabolism in the regulation of T cell function. *Trends Immunol.* 2015;36:81–91.