

Article

Metabolism-Related Gene Pairs to Predict the Clinical Outcome and Molecular Characteristics of Early Hepatocellular Carcinoma

Junling Wu ^{1,†}, Zeman Lin ^{1,†}, Daihan Ji ¹, Zhenli Li ², Huarong Zhang ¹, Shuting Lu ¹, Shenglin Wang ¹, Xiaolong Liu ^{2,*} and Lu Ao ^{1,3,*} 

- ¹ Fujian Key Laboratory of Medical Bioinformatics, Department of Bioinformatics, School of Medical Technology and Engineering, Fujian Medical University, Fuzhou 350025, China
- ² The United Innovation of Mengchao Hepatobiliary Technology Key Laboratory of Fujian Province, Mengchao Hepatobiliary Hospital of Fujian Medical University, Fuzhou 350025, China
- ³ Key Laboratory of Ministry of Education for Gastrointestinal Cancer, School of Basic Medical Sciences, Fujian Medical University, Fuzhou 350122, China
- * Correspondence: lukey@fjmu.edu.cn (L.A.); xiaoloong.liu@gmail.com (X.L.); Tel.: +86-182-5905-6924 (L.A.); +86-133-1397-5783 (X.L.)
- † These authors contributed equally to this work.

Simple Summary: After surgery, about 60–70% of early hepatocellular carcinoma patients suffer from relapse within 5 years, hindering long-term survival. Clinical and pathologic features cannot provide an accurate evaluation. We aimed to construct a stratification model from the metabolic aspect to predict the clinical outcome and reveal the molecular characteristics of different prognostic subgroups. An individualized metabolic signature of 10 gene pairs was developed from 250 early HCCs and validated in 311 samples from different datasets. The signature stratified early HCC cases one-by-one into two risk groups with different survival rates. The molecular characteristics of the two risk groups were analyzed by multi-omics data. The relationships with proliferation, immunity, and drug benefits were summarized. The signature was further validated in 47 institutional transcriptomic HCC samples and 101 public proteomic samples.



Citation: Wu, J.; Lin, Z.; Ji, D.; Li, Z.; Zhang, H.; Lu, S.; Wang, S.; Liu, X.; Ao, L. Metabolism-Related Gene Pairs to Predict the Clinical Outcome and Molecular Characteristics of Early Hepatocellular Carcinoma. *Cancers* **2022**, *14*, 3957. <https://doi.org/10.3390/cancers14163957>

Academic Editor: David A. Geller

Received: 25 July 2022

Accepted: 13 August 2022

Published: 16 August 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Abstract: Recurrence is the main factor affecting the prognosis of early hepatocellular carcinoma (HCC), which is not accurately evaluated by clinical indicators. The metabolic heterogeneity of HCC hints at the possibility of constructing a stratification model to predict the clinical outcome. On the basis of the relative expression orderings of 2939 metabolism-related genes, an individualized signature with 10 metabolism-related gene pairs (10-GPS) was developed from 250 early HCC samples in the discovery datasets, which stratified HCC patients into the high- and low-risk subgroups with significantly different survival rates. The 10-GPS was validated in 311 public transcriptomic samples from two independent validation datasets. A nomogram that included the 10-GPS, age, gender, and stage was constructed for eventual clinical evaluation. The low-risk group was characterized by lower proliferation, higher metabolism, increased activated immune microenvironment, and lower TIDE scores, suggesting a better response to immunotherapy. The high-risk group displayed hypomethylation, higher copy number alterations, mutations, and more overexpression of immune-checkpoint genes, which might jointly lead to poor outcomes. The prognostic accuracy of the 10-GPS was further validated in 47 institutional transcriptomic samples and 101 public proteomic samples. In conclusion, the 10-GPS is a robust predictor of the clinical outcome for early HCC patients and could help evaluate prognosis and characterize molecular heterogeneity.

Keywords: metabolism-related genes; hepatocellular carcinoma; risk stratification model; prognosis; molecular characteristics

1. Introduction

Liver cancer is the third-leading cause of cancer death [1]; 90% of cases are primary hepatocellular carcinoma (HCC) [2]. Surgical resection is the first choice of treatment for early HCC patients whose liver function is well-preserved. Nevertheless, approximately 60–70% of patients suffer from relapse within 5 years. The high recurrence rate after resection has become an important factor hindering the long-term survival of patients [3]. Clinical and pathologic features, such as the tumor–node–metastasis (TNM) stage, alpha-fetoprotein (AFP) level, and abnormal thrombin, cannot provide an accurate evaluation of clinical outcomes in HCC patients [4,5]. Moreover, liver cancer is considered a metabolic disease with high heterogeneity [6,7] that is selectively advantageous for tumor growth, proliferation, and survival [8]. The distinct metabolism of HCC patients hints at the possibility of constructing a stratification model from the metabolic aspect to predict the clinical outcome.

Previously reported prognostic signatures for HCC are based on risk scores obtained from the expression levels of marker genes [6,9,10], which are susceptible to batch effects. A batch of samples needs to be collected in advance and to be normalized together for the application of such signatures, which is not in accord with clinical practice. Moreover, normalizing samples together would have an impact on the risk classification of a patient [11]. By contrast, it has been demonstrated that the relative expression orderings (REOs) of genes within samples are resistant to batch effects [12] and are robust across various platforms with different designs [13]. More notably, within-sample REOs can provide personalized judgment for a single sample without data normalization. A sample is individually classified according to the REOs of gene pairs in the signature [14–16], which is more in line with actual clinical needs.

Using 2939 metabolic genes encoding human metabolic enzymes and small-molecule transporters, an individually prognostic signature was developed from 250 stage I–II HCC patients and validated in 311 transcriptomic samples from four public datasets. Then, multivariate Cox regression analysis and nomogram establishment with clinical factors were performed. The functional and clinical characteristics, epigenomic and genomic alterations, immune infiltration, and therapeutic benefits of the two prognostic groups were analyzed. The 10-GPS was further validated in 47 institutional transcriptomic samples and 101 public proteomic samples. By relating the expressions of metabolic genes with the prognosis of early HCC, the metabolism-related signature developed in this study can individually stratify patients, which will improve their prognosis, provide biological insights, predict therapeutic benefits, and assist in clinical management.

2. Materials and Methods

2.1. Study Selection Criteria

Publications that explored the prognosis of HCC and were with sufficient and available transcriptomic data and prognostic data, such as the progression-free interval (PFI, time from primary treatment to disease recurrence), disease-free survival (DFS, time from surgery to tumor recurrence, distant metastasis, or death), and overall survival (OS, time from surgery to death) were included in this study. The survival times were defined as described in the original publications [17–20]. Only TNM stage I–II HCC samples with resection were collected and analyzed.

2.2. Data Collection and Preprocessing

In total, 197 samples from The Cancer Genome Atlas [21] (TCGA, version 26.0, <https://portal.gdc.cancer.gov/>, accessed on 27 October 2020), 53 samples in GSE116174 and 170 samples in GSE14520 from Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo/>, accessed on 27 October 2020), and 141 samples from International Cancer Genome Consortium [19] (ICGC, version 28, <https://dcc.icgc.org/projects/LIRI-JP/>, accessed on 27 October 2020) were downloaded. All datasets were renamed as HCC with the sample size, e.g., HCC197. The demographic and clinicopathologic characteristics of

samples are described in Table 1. Datasets HCC197 and HCC53 were used to identify the signature; HCC170 and HCC141 were used for independent validation. The raw count represented the gene expression value of the RNA-Seq samples. The robust multiarray average algorithm was used to process the raw mRNA expression data (.CEL) from a microarray [22]. The probe ID was matched with the Entrez gene ID. One probe that matched multiple genes was deleted. The mean expression value of a gene was taken if the gene was mapped by multiple probes.

Table 1. Description of the data used in this study.

	Discovery		Validation		Institutional Validation	
	HCC197	HCC53	HCC170	HCC141	HCC101	HCC47
Accession	TCGA	GSE116174	GSE14520	LIRI-JP	CPTAC	HRA000464
Platform	Illumina Hiseq	GPL13158	GPL3921	Illumina Hiseq	Whole Proteome	Illumina Hiseq
Survival	PFI and OS	OS	DFS and OS	OS	DFS and OS	RFS and OS
Country	Mix	China	USA	Japan	China	China
Sample size	197	53	170	141	101	47
Age						
≥60	104	19	37	114	40	22
<60	93	34	133	27	61	25
Gender						
Male	142	47	143	96	77	37
Female	55	6	27	45	24	10
TNM stage						
I	138	8	93	36	87	24
II	59	45	77	105	14	23
AFP						
>300 ng/mL	34	-	66	-	64	9
≤300 ng/mL	132	-	101	-	37	38
Cirrhosis						
Yes	-	-	153	-	71	29
No	-	-	17	-	30	17
NA						1
Viral infection						
HBV	78	38	165	-	101	21
HCV	28	0	-	-	0	
HBV/HCV	5	0	-	-	0	
NA	86	15	5	-	0	26
Histologic grade						
G1/G2	118	-	-	-	-	-
G3/G4	78	-	-	-	-	-
NA	1	-	-	-	-	-
Vascular invasion						
Yes	51	-	-	-	-	-
No	134	-	-	-	-	-
NA	12	-	-	-	-	-
Child-Pugh						
A	146	-	-	-	-	-
B/C	10	-	-	-	-	-
NA	41	-	-	-	-	-

Proteomic data and clinical information of 101 stage I–II HCC patients, denoted as HCC101 (Table 1), were downloaded from the clinical proteomic tumor analysis consortium (CPTAC) [23] database. These data were used to validate the signature at the proteomic level, which might facilitate its routine implementation in clinical settings.

2.3. Samples Collection and Data Measurement

Primary tumor tissues from 72 institutional HCC patients were collected during surgery operations at Mengchao Hepatobiliary Hospital of Fujian Medical University. Primary HCC was diagnosed by at least two experienced pathologists. No patient received treatment before surgical resection. All participants signed informed consent before enrolment. All study protocols were approved by the Institution Review Board of Mengchao Hepatobiliary Hospital of Fujian Medical University and performed under the Helsinki Declaration.

RNA extracted from 72 HCC samples was subjected to whole-transcriptome sequencing on an Illumina HiSeq $\times 10$ platform (paired-end, 150 bp) by Annoroad Gene Tech. (Beijing, China) Co., Ltd. These institutional transcriptomic data of 47 patients at stage I–II, denoted as HCC47, were used to further validate the signature, and the raw counts of the genes in the RNA-Seq profiles were analyzed. The relapse-free survival (RFS, time from surgery to disease recurrence or death) and clinical information are shown in Table 1. Institutional transcriptomic data are available in the Genome Sequence Archive (GSA) repository (GSA accession number: HRA000464). More detailed information can be found in our previous work [24].

2.4. Multi-Omics Data of TCGA Portal

The multi-omics data of samples in the HCC197 dataset from TCGA were used to investigate the distinctive epigenomic and genomic characteristics between the high-risk and the low-risk subgroups. DNA methylation profiles were directly obtained from the UCSC Xena portal [20] (<https://xenabrowser.net/datapages/>, accessed on 27 November 2020). Probes with any “NA”-masked data points designed for sequences on sex chromosomes were deleted. CpG sites within the promoter regions with zero expression in more than 80% of the samples were deleted. In total, 9936 genes mapped by 17,028 CpG sites were analyzed. Somatic mutation data and copy number variation (CNV) were obtained from the GDC Data Portal (version 27.0, <https://portal.gdc.cancer.gov/>, accessed on 27 November 2020). For CNV data, the significant regions of gain or loss were identified by GISTIC 2.0 [25].

2.5. Identification of the Prognostic Signature with Metabolism-Related Gene Pairs

A list of 2939 metabolism-related genes was integrated from a previous study [26] and the latest metabolic pathways in the Kyoto Encyclopedia of Genes and Genomes [27] (KEGG) (version 97.0, <http://www.genome.jp/kegg/kegg1.html>, accessed on 15 January 2021) (Supplementary Table S1). The univariate Cox proportional-hazards regression model was used to find the candidate genes or gene pairs that were significantly correlated with the clinical outcome of early HCC. A gene whose expression level was significantly correlated with PFI in the HCC197 dataset was defined as a prognosis-associated gene. Two prognosis-associated genes, G_1 and G_2 , with expression levels of E_1 and E_2 , respectively, were constructed into a gene pair. Then, the REO pattern of the gene pair ($E_1 < E_2$ or $E_1 > E_2$) divided the samples into two groups. Gene pairs that were significantly associated with PFI were defined as prognosis-associated gene pairs. Then, all prognosis-associated gene pairs were ranked in descending order of the concordance index (C-index). Every prognosis-associated gene pair was selected as the seed, and a forward selection procedure starting with the first gene pair was performed to obtain the optimal subset that reached the highest C-index in dataset HCC197 and obtained a significant p -value of OS in dataset HCC53. The optimal subset of gene pairs was selected as the prognostic signature. The half-voting rule was used to decide the classification: a patient was classified into the high-risk group when at least half of the gene pairs in the prognostic signature voted for high risk; otherwise, the patient was classified into the low-risk group.

2.6. Survival Analysis and Differential Analysis

The multivariate Cox proportional hazards regression model was used to compute the independent prognostic value of the prognostic signature after adjusting for clinical factors. Survival curves were calculated by the log-rank test and visualized with the Kaplan–Meier plot. A nomogram with the signature, age, gender, and stage was constructed, and the 1-, 2-, and 3-year survival probability in overall samples was predicted. EdgeR and Student's *t*-test were used to identify differentially expressed genes (DEGs) in the RNA-Seq count profiles and microarray profiles, respectively. The Wilcoxon rank-sum test was used to identify significant differential methylation (DM) sites between two risk groups. Fisher's exact test was used to identify the differential frequencies of copy number alteration and mutation. Functional enrichment analysis of DEGs between the two risk groups was performed using the Database for Annotation, Visualization, and Integrated Discovery online tool (DAVID, version 6.8, <https://david.ncifcrf.gov/>, accessed on 15 January 2021). The Benjamini and Hochberg (BH) procedure was used to control the false discovery rate (FDR).

2.7. Statistical Analysis

Fisher's exact test was performed to compare the distribution of samples classified by clinical factors or the 10-GPS. Eight representative proliferation genes obtained from the study by Désert et al. [28] were used to estimate the proliferation differences in two prognostic groups (Supplementary Table S2). The median score of the principal component analysis (PCA) of the eight genes was defined as the cut-off. A sample was assigned to the non-proliferative periportal phenotype (PP) if the risk score was greater than the cut-off value; otherwise, it was assigned to the proliferative HCC subclass. The average expression levels of 43 proliferation-associated genes were also used to calculate the proliferation scores [29] (Supplementary Table S3). Moreover, 108 metabolic pathways [30] with DEGs were used to explore the metabolic differences between the two prognostic groups. The metabolic activity was scored by gene set variation analysis (GSVA) using DEGs in each pathway, and the difference was compared by limma (Supplementary Table S4). The 29 immune-associated cell types [31] calculated by single-sample GSEA (ssGSEA) and 48 immune checkpoint genes [32] were used to estimate the different immune characteristics of the two prognostic groups. Tumor immune dysfunction and exclusion (TIDE) was calculated online (<http://tide.dfci.harvard.edu/>, accessed on 15 August 2021). A high TIDE score suggested that the patients were less likely to benefit from immune checkpoint inhibition (ICI) therapy [33]. Wilcoxon rank-sum tests were performed to compare these differences. The pRRophetic algorithm was used to evaluate the 50% inhibitory concentration (IC₅₀) values of 138 antitumor drugs (including chemotherapy and targeted therapy) and identify potentially effective agents for patients in the two groups [34]. A *p*-value < 0.05 was considered significant for all analyses in this study unless otherwise stated. All statistical analyses were performed using R 3.6.3.

3. Results

3.1. Development and Validation of the Metabolism-Related Gene Pairs for Risk Stratification in Public Transcriptomic Datasets

The general flowchart of this study is described in Figure 1A. We first compared the sample distribution classified by available clinical factors in four transcriptomic datasets. The results showed that most clinical features among datasets were significantly different, suggesting the diversity of early HCC after surgery (Table 2). We then developed the signature to predict the clinical outcome of early HCC on the basis of the metabolism-related genes. In the HCC197 dataset, 164 metabolism-related genes whose expression levels were significantly correlated with the PFI of HCC patients were identified as prognosis-associated genes (FDR < 20%). Two of the 164 prognosis-associated genes were constructed into a gene pair. In total, 766 prognosis-associated gene pairs whose REOS were significantly correlated with PFI (*p* < 0.05) were identified. Then, the 10 gene pairs with the highest

C-index values (C-index = 0.6933) in the HCC197 dataset and a significant p -value of OS in the HCC53 dataset, denoted as 10-GPS, were selected as the final risk stratification signature (Figure 1B). According to the half-voting rule, a patient was classified into the high-risk group if at least five gene pairs voted for high risk; otherwise, the patient was classified into the low-risk group. The 197 samples were separately classified into the high- and low-risk groups with 64 and 133 samples. The survival analysis showed that patients in the low-risk group had significantly longer PFI and OS than those in the high-risk group (PFI: HR = 4.35, 95% CI: 2.76–6.85, $p = 6.17 \times 10^{-12}$, C-index = 0.6933; OS: HR = 3.94, 95% CI: 2.22–6.99, $p = 4.48 \times 10^{-7}$, C-index = 0.6627) (Figure 1C). In the HCC53 dataset, 34 and 19 samples were separately stratified into the high- and low-risk groups. Similarly, the OS of the latter was significantly longer than that of the former (HR = 2.98, 95% CI: 1.00–8.86, $p = 0.039$, C-index = 0.6173, Figure 1D).

Table 2. The comparison of demographics and clinical features among datasets.

Variable	HCC197 (n = 197)	HCC53 (n = 53)	HCC170 (n = 170)	HCC141 (n = 141)	p -Value	
Age	≥60	104	19	37	114	$p < 2.2 \times 10^{-16}$
Gender	Male	142	47	143	96	$p < 5.8 \times 10^{-4}$
TNM stage	I	138	8	93	36	$p < 2.2 \times 10^{-16}$
	II	59	45	77	105	
AFP	>300 ng/mL	34	-	66	-	$p < 1.9 \times 10^{-4}$
Viral infection	HBV	78	38	165	-	$p < 2.3 \times 10^{-15}$

Note. Chi-square test was used for all variables but viral infection. Fisher's exact test was used for viral infection variable.

The 10-GPS was validated on two independent datasets that included 311 stage I–II samples. Consistent with the results observed in the discovery datasets, the survival analysis showed that the low-risk group had a significantly longer OS in the HCC141 dataset (HR = 2.93, 95% CI: 1.04–8.22, $p = 0.033$, C-index = 0.6992, Figure 1E) and significantly better DFS and OS in the HCC170 dataset (DFS: HR = 1.65, 95% CI: 1.08–2.52, $p = 0.019$, C-index = 0.5638; OS: HR = 2.41, 95% CI: 1.40–4.167, $p = 0.001$, C-index = 0.5990) (Figure 1F). Univariate Cox regression analysis showed that the 10-GPS and the stage were significantly associated with the prognosis of early HCC, and multivariate Cox regression analysis confirmed that the 10-GPS remained an independent predictor after adjusting for clinical variables, including stage, histologic grade, vascular invasion, AFP, and Child–Pugh grade of cirrhosis (Figure 1G).

In addition, the univariate and multivariate Cox regression analysis showed that stage was an important factor in a patient's prognosis. Then, all samples in the four datasets were integrated, and the population was split into stages. The results showed that both stage I and stage II patients were divided by the 10-GPS into two prognostic groups with significantly different OS. Specifically, significantly more patients were predicted to be at high risk in stage II than in stage I (Chi-squared test, adjusted $p = 5 \times 10^{-6}$, Figure 2A), indicating that stage II HCC patients were more likely to be classified into the high-risk group. Furthermore, vascular invasion is also a significant factor for HCC prognosis. Patients with and without vascular invasion in HCC197 were divided by the 10-GPS into two prognostic groups with significantly different OS (Supplementary Figure S1). No difference in vascular invasion was found between the distinct prognostic groups (Chi-squared test, adjusted $p = 0.43$). These results showed that the 10-GPS was associated with the OS of early HCC independent of stage and vascular invasion.

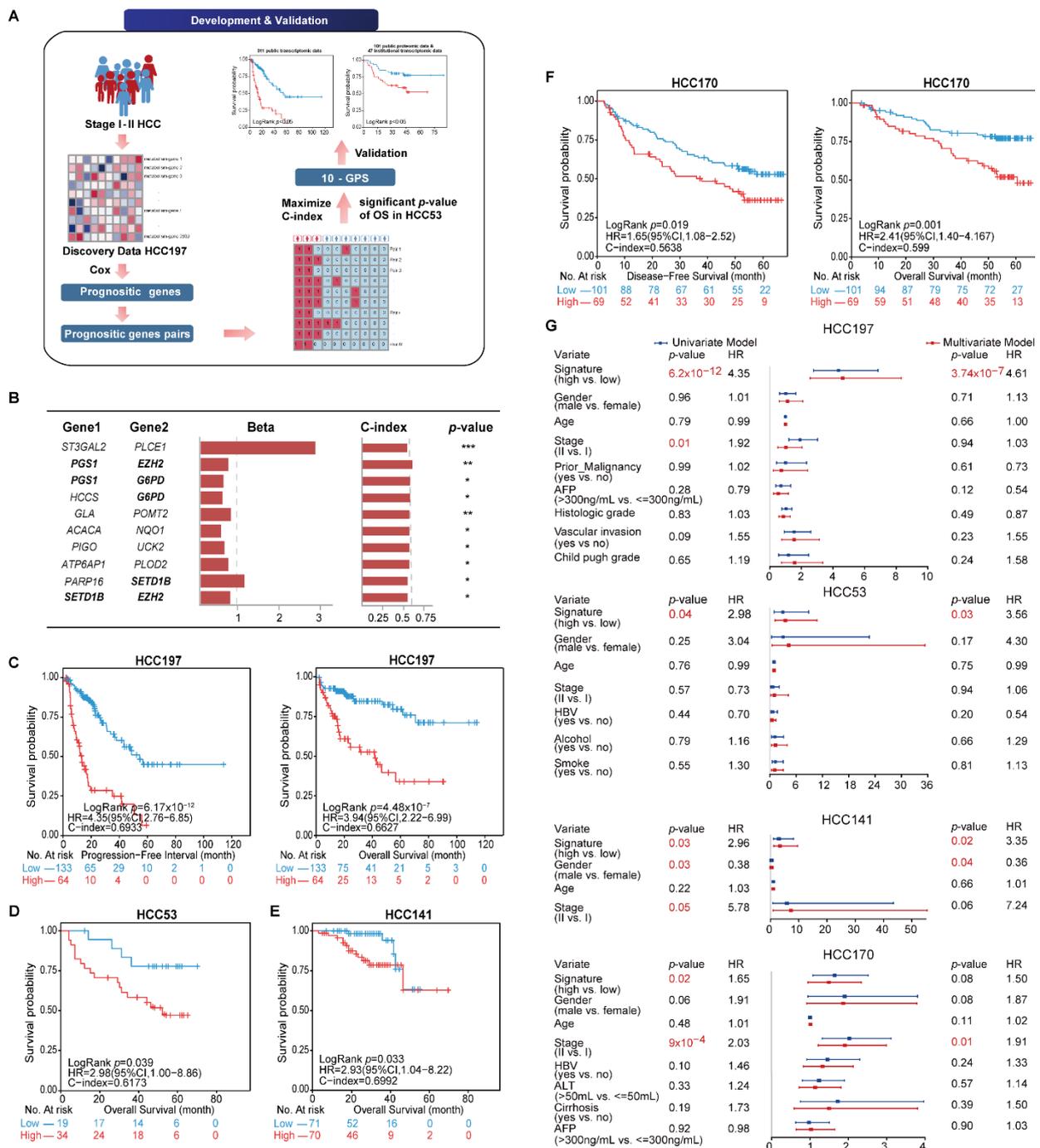


Figure 1. Identification and validation of 10-GPS in transcriptomic datasets. **(A)** The general flowchart of this study. **(B)** The 10-GPS for risk stratification. Genes shown in bold appeared twice in the signature. The univariate Cox regression model calculated the Beta and p-value. Beta was the risk coefficient of the REO of gene pair in 10-GPS, where Beta > 0 indicates that $E_1 < E_2$ is a risk factor and vice versa. *, **, and *** indicate $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. **(C,D)** The 10-GPS predicted the PFI and OS of two prognostic groups in datasets HCC197 and HCC53. According to the half-voting rule, a patient was classified into the high-risk group (red line) if at least five gene pairs voted for high risk; otherwise, the patient was classified into the low-risk group (blue line). **(E,F)** The DFS and OS of two prognostic groups in HCC141 and HCC170 were independent validation datasets. **(G)** Univariate and multivariate Cox regression analyses for the 10-GPS in the discovery and validation datasets.

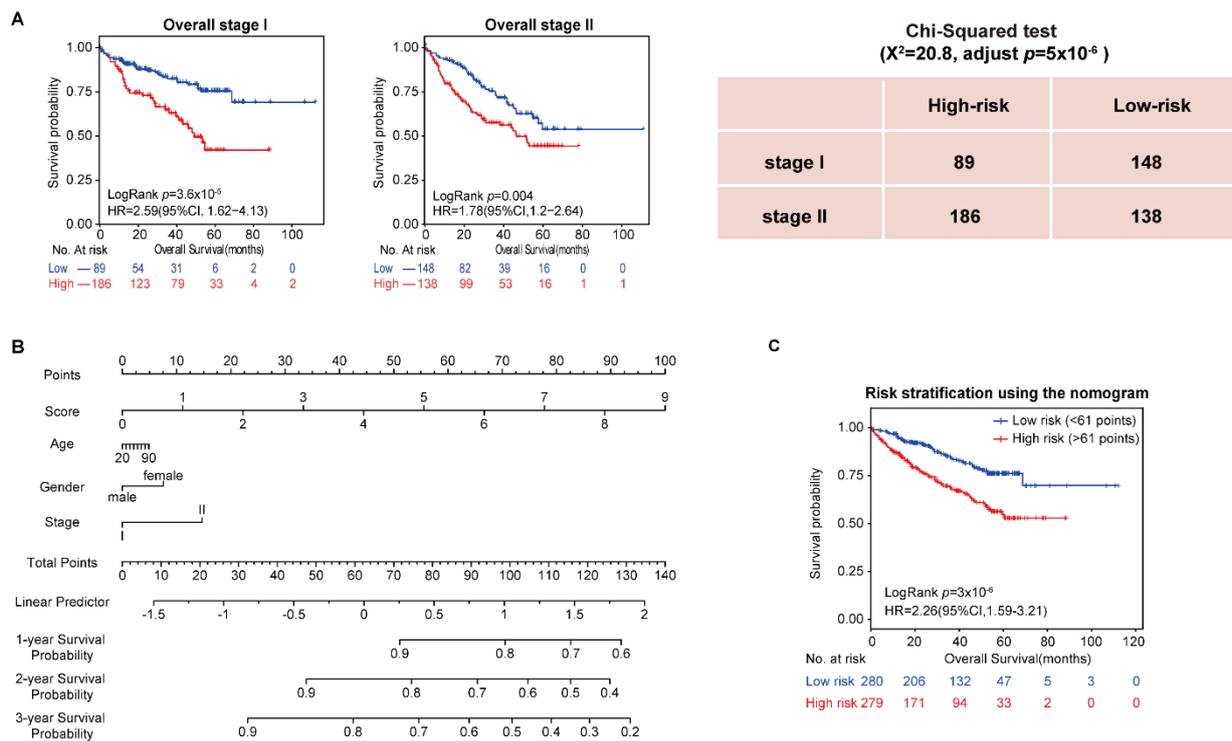


Figure 2. Analysis of 10-GPS in overall stage I–II samples. **(A)** OS of two prognostic groups in overall stage I and overall stage II samples. **(B)** Development of a composite nomogram to predict 1-year, 2-year, and 3-year survival probability. The nomogram was constructed on the basis of 10-GPS, age, gender, and stage in overall HCC samples. **(C)** The survival difference between patients divided by the median of the composite score.

Moreover, we assessed the 3 common clinical variables in the overall series of HCC patients, which included gender, age, and stage, to develop a composite prognostic predictor of the 10-GPS. The nomogram showed the contribution of each variable to predict 1-year, 2-year, and 3-year survival probability (Figure 2B). Patients divided by the median of the composite score had significantly different survival (Figure 2C).

Conclusively, these results highlighted the robustness of the 10-GPS to stratify stage I–II HCC samples from different types of data into high- and low-risk groups with significantly different prognoses.

3.2. Distinct Proliferation and Metabolism Characteristics between the Two Prognostic Groups

With 10% FDR control, 3678 and 1777 DEGs were identified between the high- and low-risk groups in the HCC197 and HCC 141 datasets (edgeR, \log_2 | Fold Change (FC) | > 1), while 20 and 1810 DEGs were identified between the two prognostic groups from the HCC53 and HCC170 datasets (Student’s *t*-test), respectively. Dataset HCC53 was excluded from the functional enrichment analysis because it had few DEGs. The overexpressed genes in the high-risk group of three datasets were all significantly enriched in cancer-related pathways, such as cell cycle and DNA replication pathways. By contrast, the underexpressed genes in the high-risk group of three datasets were all significantly enriched in metabolism-related ways, including fatty acid degradation and carbon metabolism (FDR < 5%, Figure 3A). These findings indicated that the high-risk patients with poor prognosis were considerably associated with faster proliferation ability and suppressed metabolism, which was in line with a previous report [35].

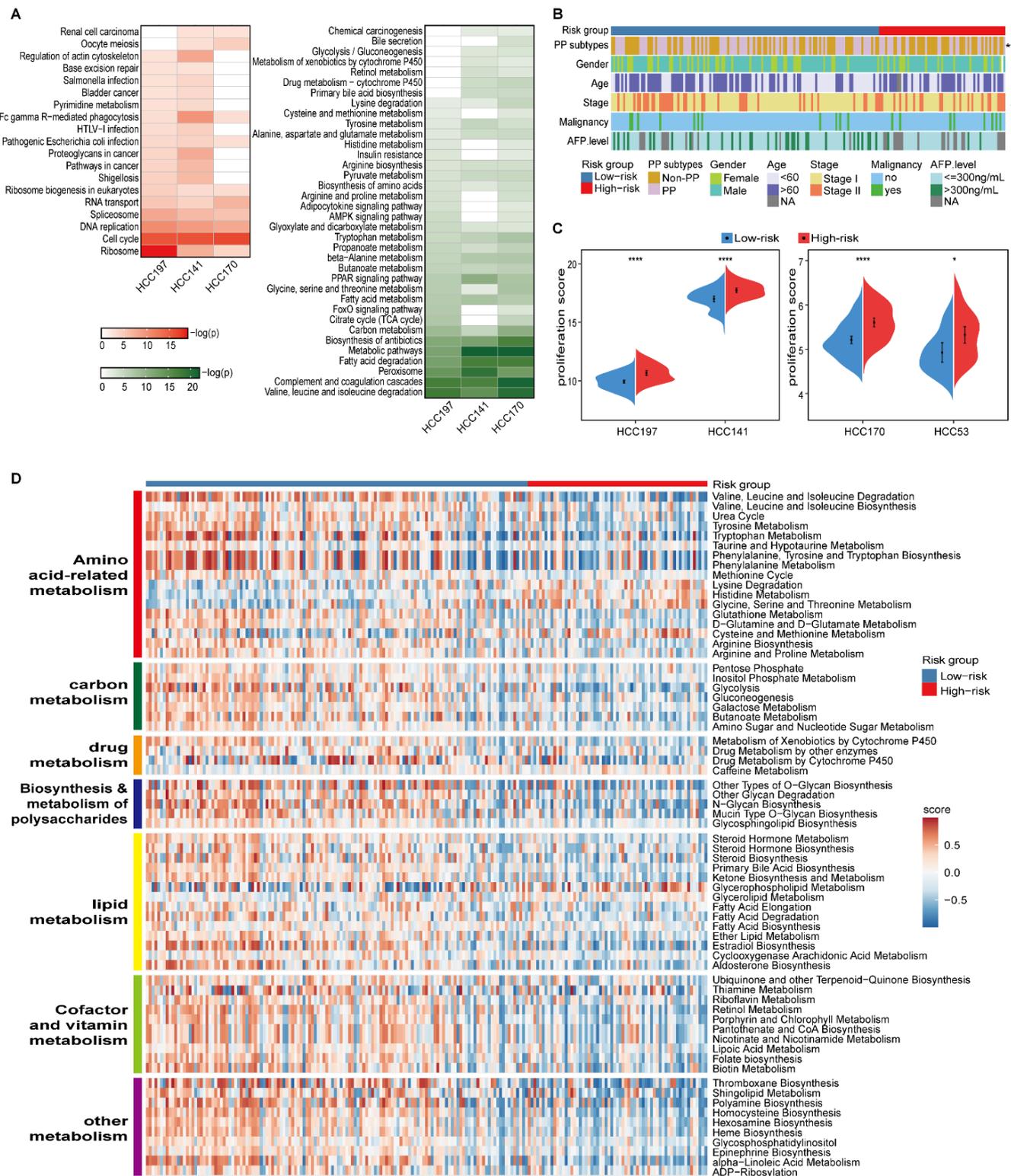


Figure 3. Distinct proliferation and metabolism characteristics of the two prognostic groups. (A) Functional enrichment analysis of DEGs between the two prognostic groups. Red and green respectively represent pathways enriched by the differentially overexpressed and underexpressed genes in the high-risk group compared with the low-risk group. The p -values were adjusted by BH (FDR < 5%). (B) The distribution of PP and non-PP subtypes, age, gender, stage, AFP level, and malignancy along with two prognostic groups. Patients were classified into the PP and non-PP subtypes according to the median value of the principal component analysis (PCA) of eight representative

proliferation genes; * $p < 0.05$, *** $p < 0.001$. (C) The split violin plots of the average expression levels of 43 proliferation-associated genes. * $p < 0.05$, **** $p < 0.0001$. (D) Heatmap of 67 differential metabolism pathways. Blue and red bars represent low- and high-risk samples, respectively. The activity scores were quantified by GSVA using DEGs in each pathway and compared by limma with $p < 0.05$.

Désert et al. reported that a non-proliferative PP subtype in HCC has the lowest potential to recur [28]. Compared with the high-risk group of HCC197, the proportion of samples classified as the PP subtype in the low-risk group was predominantly higher (44/64 versus 54/133, Fisher's exact test, $p < 0.001$, Figure 3B). Consistent results were observed in the other three datasets (Supplementary Figure S2). In addition, the proliferation ability summarized from the expression levels of 43 proliferation-relevant genes in the high-risk group was significantly higher than that in the low-risk group (Wilcoxon rank-sum test, $p < 0.05$, Figure 3C).

The metabolic characteristics between distinct prognostic groups in the HCC197 dataset were further explored on the basis of the activity of 108 metabolic pathways. The results showed that 67 metabolic pathways, which were divided into seven modules (amino acid-related metabolism, carbon metabolism, drug metabolism, biosynthesis and metabolism of polysaccharides, lipid metabolism, cofactor and vitamin metabolism, and other metabolism), were significantly different between the high- and low-risk groups (limma, FDR < 5%, Figure 3D). Furthermore, compared with the low-risk group, these metabolic pathways in the high-risk group exhibited significantly lower GSVA scores. Additionally, the clinical and pathologic characteristics, except for stage, were equally distributed between the two prognostic subgroups (Figure 3B).

Collectively, these results indicated that samples in the high-risk group exhibited a higher proliferation ability and lower metabolic activity than those in the low-risk group.

3.3. Distinct Epigenomic and Genomic Characteristics between the Two Prognostic Groups

In HCC197, 6 and 132 DM sites corresponding to 6 and 126 genes were significantly hypermethylated and hypomethylated in the high-risk group versus the low-risk group (Wilcoxon rank-sum test, FDR < 1%), respectively. Among the above 132 methylated genes, 58 genes overlapped with 3678 DEGs. Strikingly, 47 hypomethylated genes were consistently differentially overexpressed in the high-risk group, including TNFRSF11A and TP73 (Figure 4A).

The CNV analysis showed that the frequencies of copy number gain at 8q24.13 and 8q21.13 in the high-risk samples were significantly higher than the corresponding frequencies in the low-risk patients (76.56% versus 58.02% and 73.44% versus 48.85%, left of Figure 4B). There were 29 genes located in the two amplified regions, of which 27 genes were significantly overexpressed in the high-risk group (Supplementary Table S5). Overexpression of *ATAD2* and *PVT1* in 8q24.13, *RAD54B*, *LAPTM4B*, and *RRS1* in 8q21.13 were reported to be associated with tumor proliferation and progression of HCC [36–40]. Meanwhile, the frequencies of copy number loss at 13q14.2, 13q22.2, and 17p11.2 were 62.50%, 50%, and 46.88% in the high-risk patients, respectively, which were significantly higher than the corresponding frequencies of 40.46%, 27.48%, and 27.48% in the low-risk patients (right of Figure 4B). Interestingly, 11 genes located in the three deleted regions were significantly underexpressed in the high-risk group (Supplementary Table S5), including tumor suppressor gene *RB1* in 13q14.2.

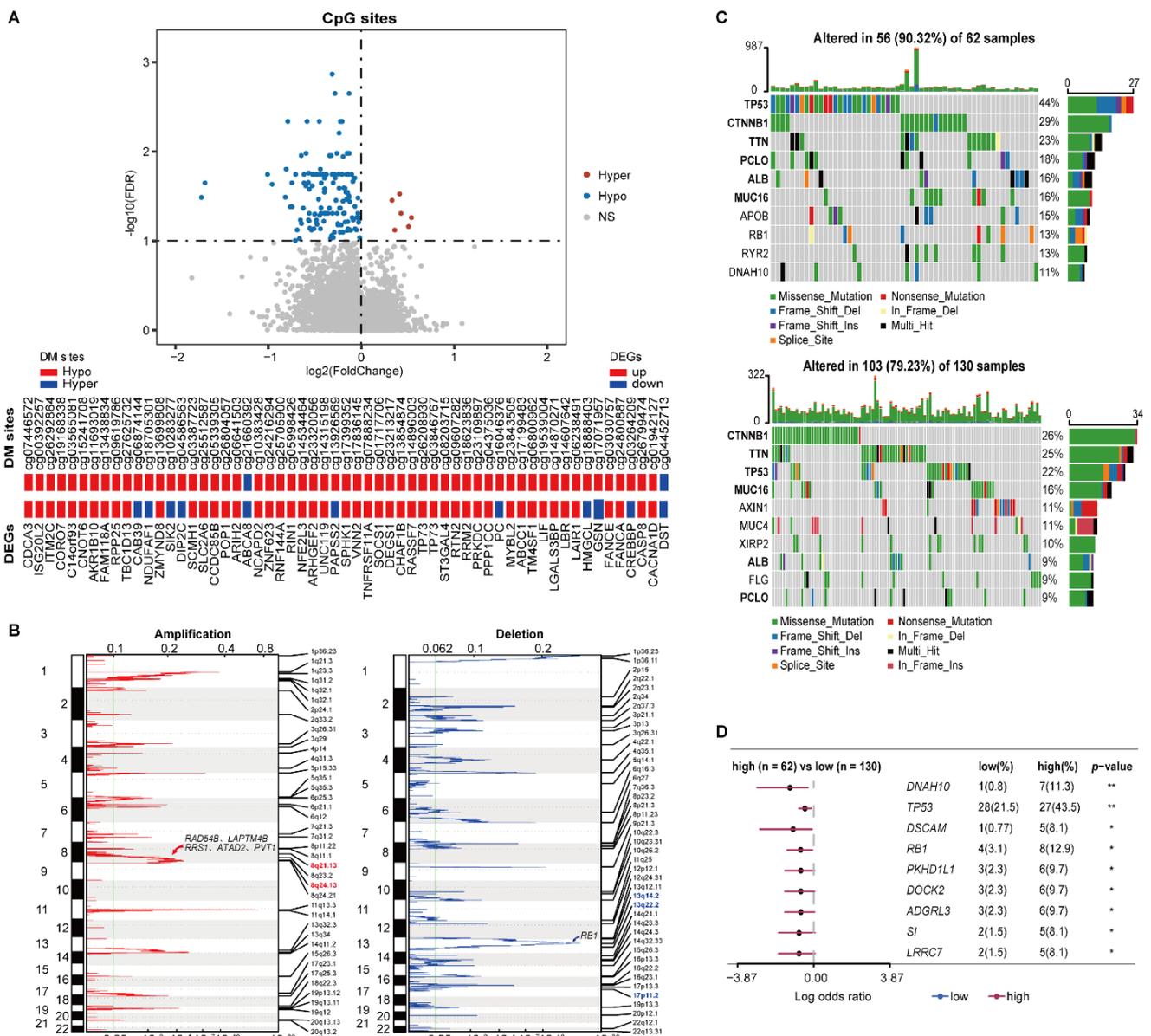


Figure 4. Multi-omics characteristics of the two prognostic groups. (A) The volcano plot of the CpG sites (the top pane) and the concordance of DM sites with DEGs (the bottom panel). Wilcoxon rank-sum test with FDR < 1% was used to identify the DM sites between two prognostic groups. (B) The CNV alterations of samples. The green line represents the cut-off point that determined the significance, q -value < 0.25. (C) The top 10 genes with the highest mutation rates in the high-risk (top) and low-risk (bottom) groups. (D) Genes with significant mutations in at least five samples from either prognostic group. Fisher’s exact test was used to calculate p -values. * p < 0.05, ** p < 0.01.

Compared with the low-risk group, a higher somatic mutation frequency was observed in the high-risk group. The top 10 genes with the highest mutation rates in the two risk subgroups are shown in Figure 4C. Notably, the driver mutation genes were not consistent between the two prognostic groups. The top six genes in the high-risk group were *TP53*, *CTNNB1*, *TNN*, *PCLO*, *ALB*, and *MUC16*, which all had varied ranks in the low-risk group. The somatic mutation on *TP53* was more prevalent in the high-risk group, which is concordant with the fact that the *TP53* mutation was positively associated with a poor prognosis of HCC [41]. Additionally, compared with the low-risk group, 75 genes had significantly higher mutation frequencies in the high-risk group (Fisher’s exact test, p < 0.05). Among them, nine genes that were mutated in at least five samples from either

prognostic group are shown in Figure 4D. Moreover, the high mutation frequency of *RB1* in the high-risk group might also contribute to its significant underexpression.

These results suggested that the epigenomic and genomic alternations exerted joint effects on the transcriptional dysregulations between the two prognostic groups, leading to poor prognoses for early HCC patients.

3.4. Distinct Immune Landscape for HCC Prognostic Groups

The immunologic landscape between different prognostic groups showed that higher scores of activated dendritic cells (aDCs), human lymphocyte antigen (HLA), MHC class I, macrophages, Th1 cells, Th2 cells, and Treg cells were observed in the high-risk group, while higher scores of B cells, mast cells, NK cells, and type II IFN response were observed in the low-risk group (Wilcoxon rank-sum test, $p < 0.05$, Figure 5A and Supplementary Figure S3). The higher scores of macrophages and Treg cells, which created an immunosuppressive microenvironment, were both correlated with a poor prognosis of HCC [42,43]. Furthermore, the higher score of type II IFN response in the low-risk group indicated an activated immune microenvironment with NK cells, which might protect against tumor development [44].

Then, we investigated the relationship between immune-checkpoint genes and the two prognostic groups. Compared with the low-risk group, more immune-checkpoint genes were significantly overexpressed in the high-risk group (Wilcoxon rank-sum test, $p < 0.05$), including *TIM3*, *B7-H3*, *TIGIT*, *TNFSF15*, and *CD44*, which further confirmed an immunosuppressive microenvironment in the high-risk group (Figures 5B and S4).

3.5. Distinct Therapeutic Benefits for HCC Prognostic Groups

We further investigated the potential drugs for two prognostic subgroups in the HCC197 dataset. In the 138-drug screen, patients in the low-risk group were predicted to be more sensitive to 40 kinds of drugs, e.g., gefitinib, whereas patients in the high-risk group were predicted to be more sensitive to the other 42 kinds of drugs, e.g., gemcitabine (Figure 5C). It is well known that gefitinib is effective against certain types of cancer. Furthermore, we observed that *EGFR*, the target gene of gefitinib, was significantly overexpressed in the low-risk group (edgeR, \log_2 | Fold Change (FC) | > 1 , FDR $< 10\%$). These results were consistent with those of a previous study, which indicated that high expression of *EGFR* was associated with a good response to gefitinib in non-small cell lung cancer [45]. Our results might provide new strategies for therapy in HCC.

Furthermore, a lower TIDE score was observed in patients in the low-risk group (Figure 5D), indicating that they might have a better response to immunotherapy than patients in the high-risk group.

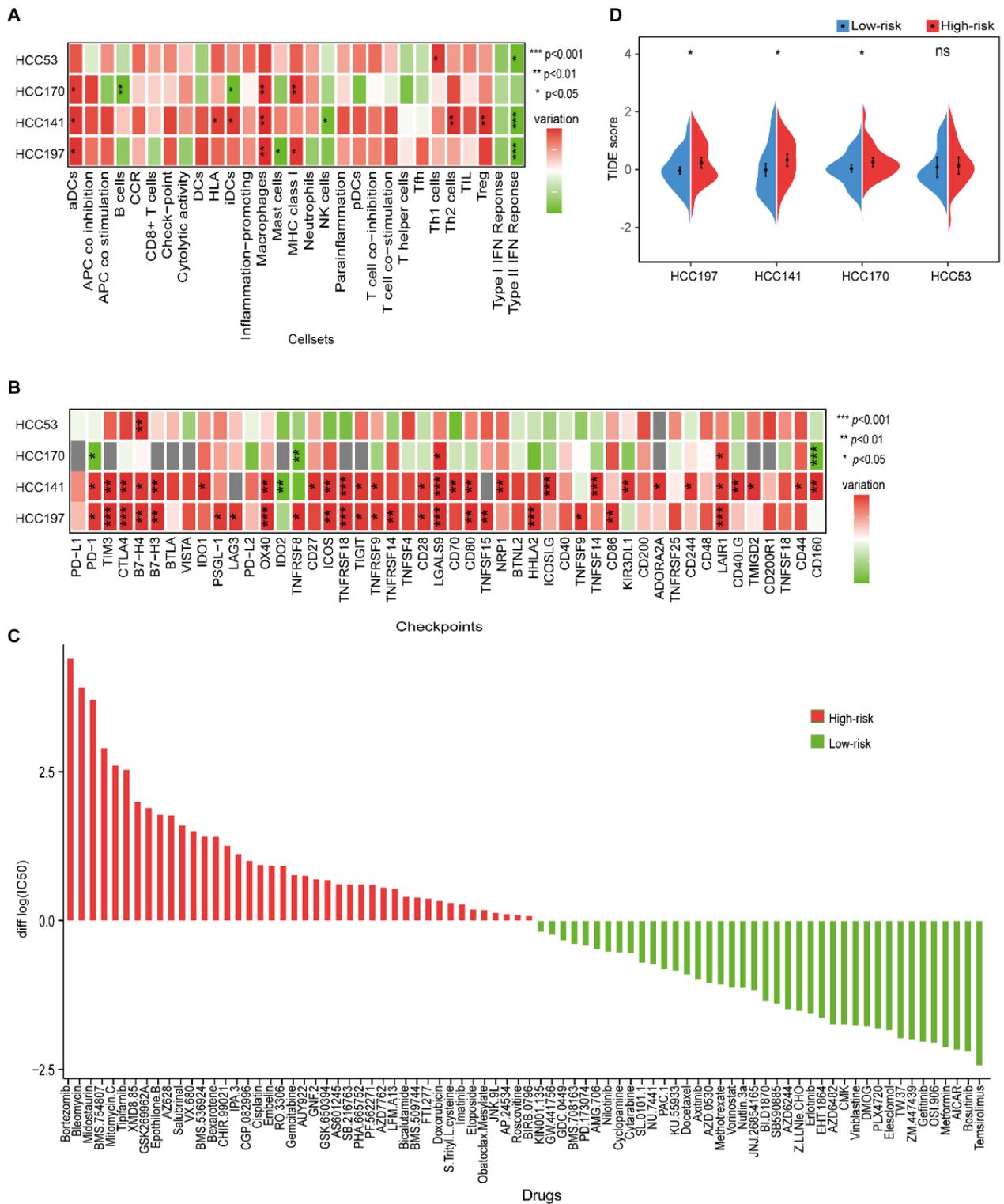


Figure 5. The immune characteristics and therapeutic benefits of two prognostic groups. **(A)** Heatmap of the levels of immune cell infiltrates in four datasets. The differences in ssGSEA scores between the high- and low-risk groups were calculated using the Wilcoxon rank-sum test. **(B)** The expression levels of 48 immune-checkpoint genes between the high- and low-risk groups in four datasets. Red and green represent a higher expression in the high- and low-risk groups, respectively. *, **, and *** indicate $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively. **(C)** Estimated drug sensitivity for patients in the high- and low-risk groups. The horizontal axis represents drugs, and the vertical axis represents the difference in the median of $\log_2 |IC_{50}|$ values between the high- and low-risk groups. Red and green

represent the drugs that might be more sensitive in the high- and low-risk groups, respectively. (D) Split violin plot of the TIDE scores for patients in the high- and low-risk groups in the four datasets.

3.6. Validation of the 10-GPS in the Institutional Transcriptomic Data and Public Proteomic Data

The prognostic performance of the 10-GPS was further evaluated in the 47 stage I–II samples generated in our previous study [24]. A total of 29 patients were classified into the low-risk group by the 10-GPS, which was associated with moderately good RFS and significantly long OS (Figure 6A). Moreover, the 10-GPS was able to predict the RFS and OS in the 72 samples (Supplementary Figure S5). These results further demonstrate that the 10-GPS can predict the prognosis of early HCC after surgery.

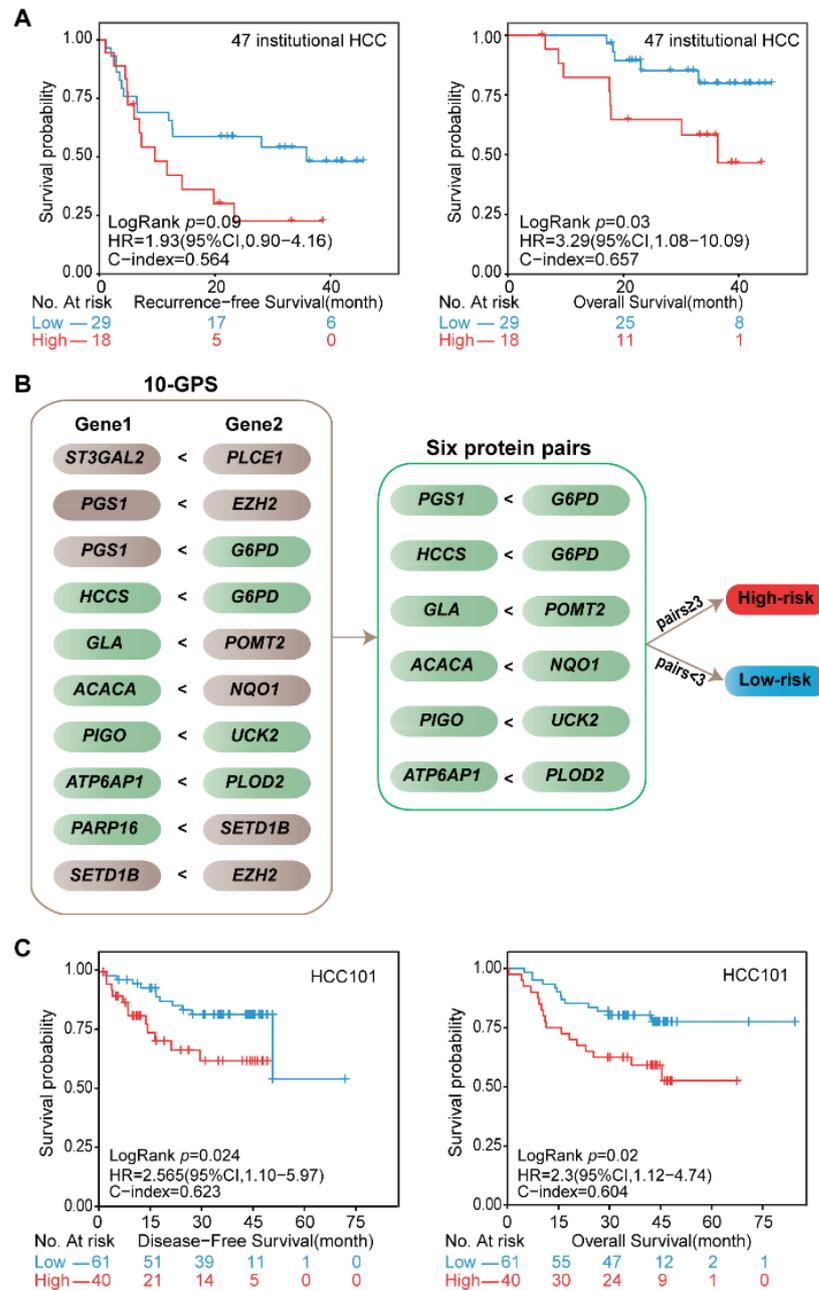


Figure 6. Validation of the signature at institutional transcriptomic data and public proteomic data. (A) Kaplan–Meier curves of the RFS and OS for the 47 institutional transcriptomic HCCd. (B) Six protein pairs in the proteomic data. Green indicates that the protein was detected in the proteomic data, and gray indicates that the gene was not detected in the proteomic data. (C) Kaplan–Meier curves of the

DFS and OS for the proteomic data of HCC101. According to the half-voting rule, a patient was classified into the high-risk group (red line) when at least three protein pairs voted for high risk; otherwise, they were classified into the low-risk group (blue line).

The proteomic data of 101 stage I–II HCC patients in the CPTAC data portal helped us to validate the 10-GPS at the proteomic level. In the proteomic data, 11 proteins encoded by the signature genes in the 10-GPS were measured and constructed into six protein pairs. Similarly, according to the half-voting rule, a patient was classified into the high-risk group when at least three protein pairs voted for high risk; otherwise, they were classified into the low-risk group (Figure 6B). The six protein pairs divided 40 patients into the high-risk group and 61 patients into the low-risk group, and the two groups had significantly different DFS (HR = 2.56, 95% CI: 1.1–5.97, $p = 0.024$, C-index = 0.623) and OS (HR = 2.30, 95% CI: 1.12–4.74, $p = 0.02$, C-index = 0.604) (Figure 6C). The results suggest that the 10-GPS can perform well in the proteomic data, which might facilitate its routine implementation in clinical settings.

4. Discussion

In this study, an individual prognostic signature consisting of 10 gene pairs with 16 metabolism-related genes, named the 10-GPS, was developed to predict clinical outcomes of early HCC. The signature was validated in multiple independent transcriptomic and proteomic datasets. Patients in the high-risk group had significantly shorter survival than patients in the low-risk group. Compared with the low-risk group, the high-risk group was characterized by lower metabolic activity, higher proliferation abilities, lower immune cells infiltration, and less immunotherapy benefit. The multi-omics analysis showed that the epigenetic and genomic alternations could corporately contribute to the transcriptional differences between the two prognostic groups. The 10-GPS was further prospectively validated in 47 institutional transcriptomic samples and 101 public proteomic samples.

At present, the treatments for early HCC and advanced HCC are considerably different. Surgical resection is the first choice of treatment for early (TNM stage I–II) HCC patients whose liver function is well-preserved but not for advanced HCC with high malignancy. Nevertheless, approximately 60–70% of patients suffer from relapse within 5 years, which is the main factor for poor prognosis. Previously, Nault et al. identified a five-gene score from I–IV stage HCC samples to predict the prognosis of HCC patients after resection [10], which might result in irrelevant features in the prognosis of early HCC. In this study, we aimed to develop a risk stratification model to predict the clinical outcome of early HCC patients after surgery, and only TNM stage I–II samples with resection were collected and analyzed. Univariate and multivariate Cox regression analysis confirmed that the 10-GPS remained an independent predictor after adjusting for clinical variables, e.g., stage and histologic grade. In addition, treatment history before hepatic resection is a potential factor for the prognosis. Although 197 stage I–II samples without any treatment before surgery from TCGA were selected as the discovery dataset, the treatment histories for samples in HCC170 and HCC53 were not unknown. The influence of treatment history on the risk stratification model for early HCC prognosis needs to be further evaluated.

We also noticed that the robustness of the 10-GPS was reduced when most genes in the signature were not expressed or detected. Our previous study and the present study revealed that the gene pair signature can be robust when more than 60% of gene pairs are detected [16]. This study also showed that 60% of protein pairs that consisted of proteins encoded by the signature genes achieved good performance at the proteomic level. However, the impact of the gene pair signature needs further assessment, and the generalization of the signature needs to be further improved and optimized in clinical application.

ICI therapy is a promising approach to improve survival in several cancers, and some immune-checkpoint proteins, e.g., *CTLA-4*, *PD-1*, and *PD-L1*, are viewed as potential biomarkers of ICI response. Patients with high expression of immune checkpoints are more likely to respond to immunotherapy [46]. Nevertheless, conflicting results were

obtained in subsequent studies [47]. Zhang et al. showed that not all patients with high expression of *PD-L1* responded to anti-*PD-L1* therapy [48]. A recent study also reported that there was no relationship between *PD-L1* expression and treatment outcome in HCC [49]. The mechanism of immune checkpoints as biomarkers for immunotherapy is not fully understood. As a comprehensive index, TIDE mimics the dysfunction and exclusion of immune cells, which might serve as a more reliable alternative biomarker for predicting the ICI response [33]. In this study, the TIDE score did not indicate that patients in the high-risk group with more immune checkpoints with high expression had a better response to immunotherapy. The complex immune microenvironment and the interaction with tumor cells might explain the conflicting results. The differences in tumor immune infiltration and the immunotherapy benefit between the two groups need further investigation in bulk and single-cell sequencing data. The findings regarding the immune landscape in Figure 4 also need to be confirmed by experiments, e.g., by IHC.

We noticed that four genes, *EZH2*, *SETD1B*, *PGS1*, and *G6PD*, appeared twice in the 10-GPS, and they might play important roles in determining the clinical outcome. Specifically, *G6PD* is the rate-limiting enzyme of the pentose phosphate pathways, which is elevated in many cancers to promote tumor growth [29]. Our results showed that *G6PD* was consistently significantly overexpressed in the high-risk group in all four transcriptomic datasets (Supplementary Figure S6A). The protein level of *G6PD* was also higher in the high-risk group than in the low-risk group. The Kaplan–Meier curves showed that high expression of *G6PD* was significantly associated with short PFI, RFS, and OS times in all four transcriptomic datasets (Supplementary Figure S6B–E). Similar results regarding *G6PD* were observed in the proteomic data of HCC101 (Supplementary Figure S6F). These results are in concordance with those of a previous study [50] and suggest that *G6PD* is a potential prognostic biomarker for HCC.

5. Conclusions

In summary, this study proposes an individual prognostic signature consisting of 16 metabolism-related genes that can accurately evaluate the outcome of early HCC and characterize inter-tumor heterogeneity in HCC.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/cancers14163957/s1>, Figure S1: Analysis of 10-GPS in samples with and without vascular invasion from HCC197; Figure S2: The distribution of PP and non-PP subtypes along with two prognostic groups; Figure S3: Distinct immune cell landscape for HCC prognostic groups; Figure S4: Distinct immune checkpoint genes for HCC prognostic groups; Figure S5: Validation of the signature in 72 institutional transcriptomic data; Figure S6: Analysis of *G6PD*; Table S1: A list of 2939 metabolism-related genes used in this study; Table S2: Eight genes were included in the periportal signature; Table S3: The 43 proliferation-associated genes; Table S4: The 108 metabolism pathways; Table S5: DEGs located in the copy number variation regions.

Author Contributions: Conceptualization, L.A.; data curation, X.L.; formal analysis, J.W., Z.L. (Zeman Lin), and Z.L. (Zhenli Li); funding acquisition, X.L. and L.A.; investigation, Z.L. (Zeman Lin), H.Z., and S.L.; methodology, J.W., D.J., S.W., and L.A.; project administration, L.A.; resources, Z.L. (Zhenli Li); software, D.J.; visualization, Z.L. (Zeman Lin); writing—original draft, J.W. and Z.L. (Zeman Lin); writing—review and editing, L.A. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the National Natural Science Foundation of China (81602738), the joint funds for the innovation of science and Technology of Fujian Province (2018Y9065), the joint research program of health and education in Fujian Province (2019-WJ-32), the Natural Science Foundation of Fujian Province (2020J01600), and the Scientific Foundation of Fuzhou Municipal Health Commission (2021-S-wp1).

Institutional Review Board Statement: All participants in HRA000464 signed informed consent before sample collection. All study protocols were approved by the Institution Review Board of Mengchao Hepatobiliary Hospital of Fujian Medical University and conducted following the principles of the Declaration of Helsinki.

Informed Consent Statement: Written informed consent was obtained from the patient(s) to publish this paper.

Data Availability Statement: Publicly available datasets were used in this study. These can be found in TCGA, ICGC, CPTAC repository, and GEO at GSE14520 and GSE116174. The institutional transcriptomic data of the 72 samples are available in the GSA repository with accession number HRA000464. Access to HRA000464 is restricted because genetic data is personally identifiable. The data and clinical information will be shared upon reasonable request to the corresponding author.

Acknowledgments: We would like to thank the TCGA, GEO, ICGC, and CPTAC databases for the availability of the data.

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

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)] [[PubMed](#)]
2. Yan, B.; Bai, D.-S.; Qian, J.-J.; Zhang, C.; Jin, S.-J.; Wang, X.; Jiang, G.-Q. Differences in tumour characteristics of Hepatocellular Carcinoma between patients with and without Cirrhosis: A population-based study. *J. Cancer* **2020**, *11*, 5812–5821. [[CrossRef](#)] [[PubMed](#)]
3. Wen, T.; Hospital, M.O.W.C.; Jin, C.; Facciorusso, A.; Donadon, M.; Han, H.-S.; Mao, Y.; Dai, C.; Cheng, S.; Zhang, B.; et al. Multidisciplinary management of recurrent and metastatic hepatocellular carcinoma after resection: An international expert consensus. *Hepatobiliary Surg. Nutr.* **2018**, *7*, 353–371. [[CrossRef](#)]
4. Yang, A.; Xiao, W.; Chen, N.; Wei, X.; Huang, S.; Lin, Y.; Zhang, C.; Lin, J.; Deng, F.; Wu, C.; et al. The power of tumor sizes in predicting the survival of solitary hepatocellular carcinoma patients. *Cancer Med.* **2018**, *7*, 6040–6050. [[CrossRef](#)]
5. Ahn, J.C.; Teng, P.C.; Chen, P.J.; Posadas, E.; Tseng, H.R.; Lu, S.C.; Yang, J.D. Detection of circulating tumor cells and their implications as a novel biomarker for diagnosis, prognostication, and therapeutic monitoring in hepatocellular carcinoma. *Hepatology* **2021**, *73*, 422–436. [[CrossRef](#)] [[PubMed](#)]
6. Yang, C.; Huang, X.; Liu, Z.; Qin, W.; Wang, C. Metabolism-associated molecular classification of hepatocellular carcinoma. *Mol. Oncol.* **2020**, *14*, 896–913. [[CrossRef](#)] [[PubMed](#)]
7. Orabi, D.; Berger, N.A.; Brown, J.M. Abnormal Metabolism in the Progression of Nonalcoholic Fatty Liver Disease to Hepatocellular Carcinoma: Mechanistic Insights to Chemoprevention. *Cancers* **2021**, *13*, 3473. [[CrossRef](#)]
8. Boroughs, L.K.; DeBerardinis, R.J. Metabolic pathways promoting cancer cell survival and growth. *Nat. Cell Biol.* **2015**, *17*, 351–359. [[CrossRef](#)]
9. Roayaie, S.; Blume, I.N.; Thung, S.N.; Guido, M.; Fiel, M.; Hiotis, S.; Labow, D.M.; Llovet, J.M.; Schwartz, M.E. A System of Classifying Microvascular Invasion to Predict Outcome After Resection in Patients With Hepatocellular Carcinoma. *Gastroenterology* **2009**, *137*, 850–855. [[CrossRef](#)]
10. Nault, J.C.; De Reynies, A.; Villanueva, A.; Calderaro, J.; Rebouissou, S.; Couchy, G.; Decaens, T.; Franco, D.; Imbeaud, S.; Rousseau, F.; et al. A hepato-cellular carcinoma 5-gene score associated with survival of patients after liver resection. *Gastroenterology* **2013**, *145*, 176–187. [[CrossRef](#)]
11. Qi, L.; Chen, L.; Li, Y.; Qin, Y.; Pan, R.; Zhao, W.; Gu, Y.; Wang, H.; Wang, R.; Chen, X.; et al. Critical limitations of prognostic signatures based on risk scores summarized from gene expression levels: A case study for resected stage I non-small-cell lung cancer. *Brief. Bioinform.* **2016**, *17*, 233–242. [[CrossRef](#)] [[PubMed](#)]
12. Geman, D.; D’Avignon, C.; Naiman, D.Q.; Winslow, R.L. Classifying Gene Expression Profiles from Pairwise mRNA Comparisons. *Stat. Appl. Genet. Mol. Biol.* **2004**, *3*, 1–19. [[CrossRef](#)] [[PubMed](#)]
13. Guan, Q.; Chen, R.; Yan, H.; Cai, H.; Guo, Y.; Li, M.; Li, X.; Tong, M.; Ao, L.; Li, H.; et al. Differential expression analysis for individual cancer samples based on robust within-sample relative gene expression orderings across multiple profiling platforms. *Oncotarget* **2016**, *7*, 68909–68920. [[CrossRef](#)]
14. Wang, H.; Cai, H.; Ao, L.; Yan, H.; Zhao, W.; Qi, L.; Gu, Y.; Guo, Z. Individualized identification of disease-associated pathways with disrupted coordination of gene expression. *Brief. Bioinform.* **2016**, *17*, 78–87. [[CrossRef](#)]
15. Ao, L.; Zhang, Z.; Guan, Q.; Guo, Y.; Guo, Y.; Zhang, J.; Lv, X.; Huang, H.; Zhang, H.; Wang, X.; et al. A qualitative signature for early diagnosis of hepatocellular carcinoma based on relative expression orderings. *Liver Int.* **2018**, *38*, 1812–1819. [[CrossRef](#)] [[PubMed](#)]
16. Huang, H.; Zou, Y.; Zhang, H.; Li, X.; Li, Y.; Deng, X.; Sun, H.; Guo, Z.; Ao, L. A qualitative transcriptional prognostic signature for patients with stage I-II pancreatic ductal adenocarcinoma. *Transl. Res.* **2020**, *219*, 30–44. [[CrossRef](#)]

17. Roessler, S.; Jia, H.-L.; Budhu, A.; Forgues, M.; Ye, Q.-H.; Lee, J.-S.; Thorgeirsson, S.S.; Sun, Z.; Tang, Z.-Y.; Qin, L.-X.; et al. A Unique Metastasis Gene Signature Enables Prediction of Tumor Relapse in Early-Stage Hepatocellular Carcinoma Patients. *Cancer Res.* **2010**, *70*, 10202–10212. [[CrossRef](#)]
18. Liu, J.; Lichtenberg, T.M.; Hoadley, K.A.; Poisson, L.M.; Lazar, A.J.; Cherniack, A.D.; Kovatich, A.J.; Benz, C.C.; Levine, D.A.; Lee, A.V.; et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell* **2018**, *173*, 400–416.e11. [[CrossRef](#)]
19. Zhang, J.; Bajari, R.; Andric, D.; Gerthoffert, F.; Lepsa, A.; Nahal-Bose, H.; Stein, L.D.; Ferretti, V. The International Cancer Genome Consortium Data Portal. *Nat. Biotechnol.* **2019**, *37*, 367–369. [[CrossRef](#)]
20. Goldman, M.J.; Craft, B.; Hastie, M.; Repčeka, K.; McDade, F.; Kamath, A.; Banerjee, A.; Luo, Y.; Rogers, D.; Brooks, A.N.; et al. Visualizing and interpreting cancer genomics data via the Xena platform. *Nat. Biotechnol.* **2020**, *38*, 675–678. [[CrossRef](#)]
21. Blum, A.; Wang, P.; Zenklusen, J.C. SnapShot: TCGA-Analyzed Tumors. *Cell* **2018**, *173*, 530. [[CrossRef](#)] [[PubMed](#)]
22. Irizarry, R.A.; Hobbs, B.; Collin, F.; Beazer-Barclay, Y.D.; Antonellis, K.J.; Scherf, U.; Speed, T. Exploration, normalization, and summaries of high density oligonucleotide array probe level data. *Biostatistics* **2003**, *4*, 249–264. [[CrossRef](#)] [[PubMed](#)]
23. Edwards, N.J.; Oberti, M.; Thangudu, R.R.; Cai, S.; McGarvey, P.B.; Jacob, S.; Madhavan, S.; Ketchum, K.A. The CPTAC Data Portal: A Resource for Cancer Proteomics Research. *J. Proteome Res.* **2015**, *14*, 2707–2713. [[CrossRef](#)]
24. Li, Z.; Chen, G.; Cai, Z.; Dong, X.; He, L.; Qiu, L.; Zeng, Y.; Liu, X.; Liu, J. Profiling of hepatocellular carcinoma neoantigens reveals immune microenvironment and clonal evolution related patterns. *Chin. J. Cancer Res.* **2021**, *33*, 364–378. [[CrossRef](#)]
25. Mermel, C.H.; Schumacher, S.E.; Hill, B.; Meyerson, M.L.; Beroukhi, R.; Getz, G. GISTIC2.0 facilitates sensitive and confident localization of the targets of focal somatic copy-number alteration in human cancers. *Genome Biol.* **2011**, *12*, R41. [[CrossRef](#)]
26. Possemato, R.; Marks, K.M.; Shaul, Y.D.; Pacold, M.E.; Kim, D.; Birsoy, K.; Sethumadhavan, S.; Woo, H.-K.; Jang, H.G.; Jha, A.K.; et al. Functional genomics reveal that the serine synthesis pathway is essential in breast cancer. *Nature* **2011**, *476*, 346–350. [[CrossRef](#)] [[PubMed](#)]
27. Huang, D.W.; Sherman, B.T.; Lempicki, R.A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **2009**, *4*, 44–57. [[CrossRef](#)]
28. Désert, R.; Rohart, F.; Canal, F.; Sicard, M.; Desille, M.; Renaud, S.; Turlin, B.; Bellaud, P.; Perret, C.; Clément, B.; et al. Human hepatocellular carcinomas with a periportal phenotype have the lowest potential for early recurrence after curative resection. *Hepatology* **2017**, *66*, 1502–1518. [[CrossRef](#)] [[PubMed](#)]
29. Whitfield, M.L.; George, L.K.; Grant, G.; Perou, C. Common markers of proliferation. *Nat. Rev. Cancer* **2006**, *6*, 99–106. [[CrossRef](#)]
30. Rosario, S.R.; Long, M.D.; Affronti, H.C.; Rowsam, A.M.; Eng, K.H.; Smiraglia, D.J. Pan-cancer analysis of transcriptional metabolic dysregulation using The Cancer Genome Atlas. *Nat. Commun.* **2018**, *9*, 5330. [[CrossRef](#)]
31. He, Y.; Jiang, Z.; Chen, C.; Wang, X. Classification of triple-negative breast cancers based on Immunogenomic profiling. *J. Exp. Clin. Cancer Res.* **2018**, *37*, 327. [[CrossRef](#)] [[PubMed](#)]
32. Wu, J.; Li, L.; Zhang, H.; Zhao, Y.; Zhang, H.; Wu, S.; Xu, B. A risk model developed based on tumor microenvironment predicts overall survival and associates with tumor immunity of patients with lung adenocarcinoma. *Oncogene* **2021**, *40*, 4413–4424. [[CrossRef](#)]
33. Jiang, P.; Gu, S.; Pan, D.; Fu, J.; Sahu, A.; Hu, X.; Li, Z.; Traugh, N.; Bu, X.; Li, B.; et al. Signatures of T cell dysfunction and exclusion predict cancer immunotherapy response. *Nat. Med.* **2018**, *24*, 1550–1558. [[CrossRef](#)]
34. Geeleher, P.; Cox, N.; Huang, R.S. pRRophetic: An R Package for Prediction of Clinical Chemotherapeutic Response from Tumor Gene Expression Levels. *PLoS ONE* **2014**, *9*, e107468. [[CrossRef](#)] [[PubMed](#)]
35. Gao, Q.; Zhu, H.; Dong, L.; Shi, W.; Chen, R.; Song, Z.; Huang, C.; Li, J.; Dong, X.; Zhou, Y.; et al. Integrated Proteogenomic Characterization of HBV-Related Hepatocellular Carcinoma. *Cell* **2019**, *179*, 561–577.e22. [[CrossRef](#)]
36. Hwang, H.W.; Ha, S.Y.; Bang, H.; Park, C.-K. ATAD2 as a Poor Prognostic Marker for Hepatocellular Carcinoma after Curative Resection. *Cancer Res. Treat.* **2015**, *47*, 853–861. [[CrossRef](#)] [[PubMed](#)]
37. Zhang, Y.; Wen, D.-Y.; Zhang, R.; Huang, J.-C.; Lin, P.; Ren, F.-H.; Wang, X.; He, Y.; Yang, H.; Chen, G.; et al. A Preliminary Investigation of PVT1 on the Effect and Mechanisms of Hepatocellular Carcinoma: Evidence from Clinical Data, a Meta-Analysis of 840 Cases, and In Vivo Validation. *Cell. Physiol. Biochem.* **2018**, *47*, 2216–2232. [[CrossRef](#)] [[PubMed](#)]
38. Feng, S.; Liu, J.; Hailiang, L.; Wen, J.; Zhao, Y.; Li, X.; Lu, G.; Gao, P.; Zeng, X. Amplification of RAD54B promotes progression of hepatocellular carcinoma via activating the Wnt/beta-catenin signaling. *Transl. Oncol.* **2021**, *14*, 101124. [[CrossRef](#)]
39. Wang, F.; Wu, H.; Zhang, S.; Lu, J.; Lu, Y.; Zhan, P.; Fang, Q.; Wang, F.; Zhang, X.; Xie, C.; et al. LAPTM4B facilitates tumor growth and induces autophagy in hepatocellular carcinoma. *Cancer Manag. Res.* **2019**, *11*, 2485–2497. [[CrossRef](#)]
40. Yan, X.; Wu, S.; Liu, Q.; Zhang, J. RRS1 Promotes Retinoblastoma Cell Proliferation and Invasion via Activating the AKT/mTOR Signaling Pathway. *BioMed Res. Int.* **2020**, *2020*, 1–10. [[CrossRef](#)]
41. Yang, C.; Huang, X.; Li, Y.; Chen, J.; Lv, Y.; Dai, S. Prognosis and personalized treatment prediction in TP53-mutant hepatocellular carcinoma: An in silico strategy towards precision oncology. *Brief. Bioinform.* **2020**, *22*, bbaa164. [[CrossRef](#)] [[PubMed](#)]
42. Wu, Q.; Zhou, W.; Yin, S.; Zhou, Y.; Chen, T.; Qian, J.; Su, R.; Hong, L.; Lu, H.; Zhang, F.; et al. Blocking Triggering Receptor Expressed on Myeloid Cells-1-Positive Tumor-Associated Macrophages Induced by Hypoxia Reverses Immuno-suppression and Anti-Programmed Cell Death Ligand 1 Resistance in Liver Cancer. *Hepatology* **2019**, *70*, 198–214. [[CrossRef](#)] [[PubMed](#)]

43. Yang, P.; Li, Q.J.; Feng, Y.; Zhang, Y.; Markowitz, G.J.; Ning, S.; Deng, Y.; Zhao, J.; Jiang, S.; Yuan, Y.; et al. TGF-beta-miR-34a-CCL22 signaling-induced Treg cell recruitment promotes venous metastases of HBV-positive hepato-cellular carcinoma. *Cancer Cell* **2012**, *22*, 291–303. [[CrossRef](#)] [[PubMed](#)]
44. Ikeda, H.; Old, L.J.; Schreiber, R.D. The roles of IFN gamma in protection against tumor development and cancer immunoediting. *Cytokine Growth Factor Rev.* **2002**, *13*, 95–109. [[CrossRef](#)]
45. Ono, M.; Kuwano, M. Molecular Mechanisms of Epidermal Growth Factor Receptor (EGFR) Activation and Response to Gefitinib and Other EGFR-Targeting Drugs. *Clin. Cancer Res.* **2006**, *12*, 7242–7251. [[CrossRef](#)]
46. Pare, L.; Pascual, T.; Segui, E.; Teixido, C.; Gonzalez-Cao, M.; Galvan, P.; Rodriguez, A.; Gonzalez, B.; Cuatrecasas, M.; Pineda, E.; et al. Association between PD1 mRNA and response to anti-PD1 monotherapy across multiple cancer types. *Ann. Oncol.* **2018**, *29*, 2121–2128. [[CrossRef](#)]
47. Nishino, M.; Ramaiya, N.H.; Hatabu, H.; Hodi, F.S. Monitoring immune-checkpoint blockade: Response evaluation and biomarker development. *Nat. Rev. Clin. Oncol.* **2017**, *14*, 655–668. [[CrossRef](#)]
48. Zhang, Y.; Chen, H.; Mo, H.; Hu, X.; Gao, R.; Zhao, Y.; Liu, B.; Niu, L.; Sun, X.; Yu, X.; et al. Single-cell analyses reveal key immune cell subsets associated with response to PD-L1 blockade in triple-negative breast cancer. *Cancer Cell* **2021**, *39*, 1578–1593.e8. [[CrossRef](#)]
49. El-Khoueiry, A.B.; Sangro, B.; Yau, T.; Crocenzi, T.S.; Kudo, M.; Hsu, C.; Kim, T.-Y.; Choo, S.-P.; Trojan, J.; Welling, T.H., 3rd; et al. Nivolumab in patients with advanced hepatocellular carcinoma (CheckMate 040): An open-label, non-comparative, phase 1/2 dose escalation and expansion trial. *Lancet* **2017**, *389*, 2492–2502. [[CrossRef](#)]
50. Li, M.; He, X.; Guo, W.; Yu, H.; Zhang, S.; Wang, N.; Liu, G.; Sa, R.; Shen, X.; Jiang, Y.; et al. Aldolase B suppresses hepatocellular carcinogenesis by inhibiting G6PD and pentose phosphate pathways. *Nat. Cancer* **2020**, *1*, 735–747. [[CrossRef](#)]