

Research Article

Identification of Hypoxia-Related Differentially Expressed Genes and Construction of the Clinical Prognostic Predictor in Hepatocellular Carcinoma by Bioinformatic Analysis

Hao Guo,¹ Jing Zhou ,² Yanjun Zhang,³ Zhi Wang,¹ Likun Liu,² and Weiwei Zhang¹

¹Department of Anesthesiology, Shanxi Provincial People's Hospital, Taiyuan, Shanxi 030000, China

²Department of Oncology, Shanxi Province Academy of Traditional Chinese Medicine, Shanxi Province Hospital of Traditional Chinese Medicine, Taiyuan, Shanxi 030012, China

³Department of Anesthesiology, Tianjin Children's Hospital, Tianjin 300074, China

Correspondence should be addressed to Jing Zhou; zhoujingsx07@163.com

Received 21 June 2020; Revised 13 December 2020; Accepted 21 January 2021; Published 8 February 2021

Academic Editor: Aparup Das

Copyright © 2021 Hao Guo et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Background. Hypoxia closely relates to malignant progression and appears to be prognostic for outcome in hepatocellular carcinoma (HCC). Our research is aimed at mining the hypoxic-related genes (HRGs) and constructing a prognostic predictor (PP) model on clinical prognosis in HCC patients. **Methods.** RNA-sequencing data about HRGs and clinical data of patients with HCC were obtained from The Cancer Genome Atlas (TCGA) database portal. Differentially expressed HRGs between HCC and para-carcinoma tissue samples were obtained by applying the Wilcox analysis in R statistical software. Gene Ontology and Kyoto Encyclopedia of Genes and Genomes were used for gene functional enrichment analyses. Then, the patients who were asked to follow up for at least one month were enrolled in the following study. Cox proportional risk regression model was applied to obtain key HRGs which related to overall survival (OS) in HCC. PP was constructed and defined, and the accuracy of PP was validated by constructing the signature in a training set and validation set. Connectivity map (CMap) was used to find potential drugs, and gene set cancer analysis (GSCA) was also performed to explore the underlying molecular mechanisms. **Results.** Thirty-seven differentially expressed HRGs were obtained. It contained 28 upregulated and 9 downregulated genes. After the univariate Cox regression model analysis, we obtained 27 prognosis-related HRGs. Of these, 25 genes were risk factors for cancer, and 2 genes were protective factors. The PP was composed by 12 key genes (HDLBP, SAP30, PFKP, DPYSL4, SLC2A1, HMOX1, PGK1, ERO1A, LDHA, ENO2, SLC6A6, and TPI1). GSCA results showed the overall activity of these 12 key genes in 10 cancer-related pathways. Besides, CMap identified deferoramine, crotamiton, talampicillin, and lycorine might have effects with HCC. **Conclusions.** This study firstly reported 12 prognostic HRGs and constructed the model of the PP. This comprehensive research of multiple databases helps us gain insight into the biological properties of HCC and provides deferoramine, crotamiton, talampicillin, and lycorine as potential drugs to fight against HCC.

1. Introduction

Hypoxia is a hallmark of the tumor microenvironment and contributes to tumor malignant phenotypes. It presents in the majority of tumors, and the distribution of hypoxic regions in tumors is heterogeneous [1]. Hypoxia has been considered as a vital sign of poor prognosis in almost all solid tumors, including the hepatocellular carcinoma (HCC) [2]. Hypoxia can lead to changes in the proteome and/or genome.

The mechanism may be related to cells' continued access to nutrients, escape from unsatisfactory microenvironments, and promote unrestricted growth to accelerate tumor progression [3]. Tumor cells in hypoxic regions are closely associated with treatment resistance, local recurrence, and distant metastasis [4]. All these can be regarded as important factors for poor prognosis of patients.

Liver cancer is the fourth leading cause of cancer death in 2018. It kills about 782,000 people annually [5, 6]. It has been

predicted that liver cancer may become the sixth most commonly diagnosed cancer in the future. HCC comprises 75%–85% of primary cancer of the liver. Unfortunately, the exact mechanism and pathways leading to HCC development are still unclear explanation [7]. In particular, the absence of indicators is related to the prognosis of HCC. The 5-year survival rate is still very low [8, 9]. Therefore, finding the key molecular biomarkers focused on hypoxia which is related to the occurrence and development of HCC patients has certain reference value for reliable estimation of the deterioration of HCC and may be an effective measure to against HCC.

In recent years, progress using high-throughput sequencing technologies has greatly expanded the research of cancer genome. Here, we compared the expression profile of hypoxic-related genes (HRGs) between HCC tissue samples and para-carcinoma tissue samples which were collected from The Cancer Genome Atlas (TCGA) database [10]. Furthermore, in order to clarify the correlation between the HRGs and clinical outcomes in HCC patients, the prognostic predictor (PP) was developed and validated as an independent indicator to assess the overall survival (OS) of patients. We also tried to find potential molecular drugs to fight against HCC. These findings may provide new research targets and therapeutic strategies for HCC patients and provide a reliable theoretical basis for judging treatment outcomes and assessing prognosis.

2. Materials and Methods

2.1. Data Acquisition. The Molecular Signatures Database (MSigDB, <http://software.broadinstitute.org/gsea/msigdb/index.jsp>) is a collection of annotated gene sets for GSEA software. A variety of HRGs were extracted from the Hallmark hypoxic gene sets in the MSigDB [11, 12], which contains a list of 200 HRGs involved in hypoxic regulatory pathways reported in the literature. RNA-sequencing data about HRGs and clinical data of patients with HCC were obtained from the TCGA data portal. The dataset consists of 374 HCC tissue samples and 50 para-carcinoma tissue samples. We performed on these samples to obtain the differentially expressed HRGs. Then, in order to build a model related to the prognosis of HCC, a total of 349 patients who were asked to follow up for at least one month were enrolled in our study. The follow-up periods of included patients were for a range of 30 days to 3675 days.

2.2. Differentially Expressed HRG Enrichment Analysis. We screened for differences in HRG expression between HCC tissue samples and para-carcinoma tissue samples by applying the Wilcox analysis in R statistical software. The significantly differentially expressed HRGs were defined as the p value is less than 0.05; the gene expression changes at least 1.5-fold. To explain the main biological properties of HRGs, the tools of Gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were used for gene functional enrichment analyses. GO and KEGG terms with a p value and q value both smaller than 0.05 were considered significant categories. To take full advantage of the well-

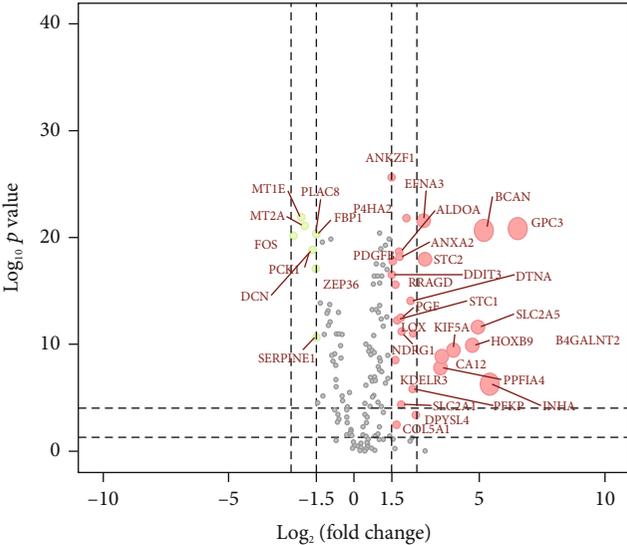
known GO and KEGG themes and highlight the most relevant GO terms associated with a given gene list, these results of gene enrichment were visualized by R language clustering package [13].

2.3. Construction of the Individualized Prognostic Predictor Model Based on HRGs. We extracted the expression profile of HRGs downloaded from TCGA in the FPKM format. Univariate Cox regression analysis was performed to screen HRGs whose expression profile was significantly correlated with OS in HCC patients from 200 hypoxic-related genes ($p < 0.01$). Multivariate Cox regression analysis was performed on these survival-related genes to remove genes that might not be independent indicators of prognosis. We got the optimal solution of the model by narrowing down the relevant genes step by step. Finally, the PP was constructed and significantly divided patients of HCC into high- and low-risk groups according to OS. Kaplan-Meier (K-M) method was used to plot the survival curve. The survival rates of the two groups were tested by log-rank test.

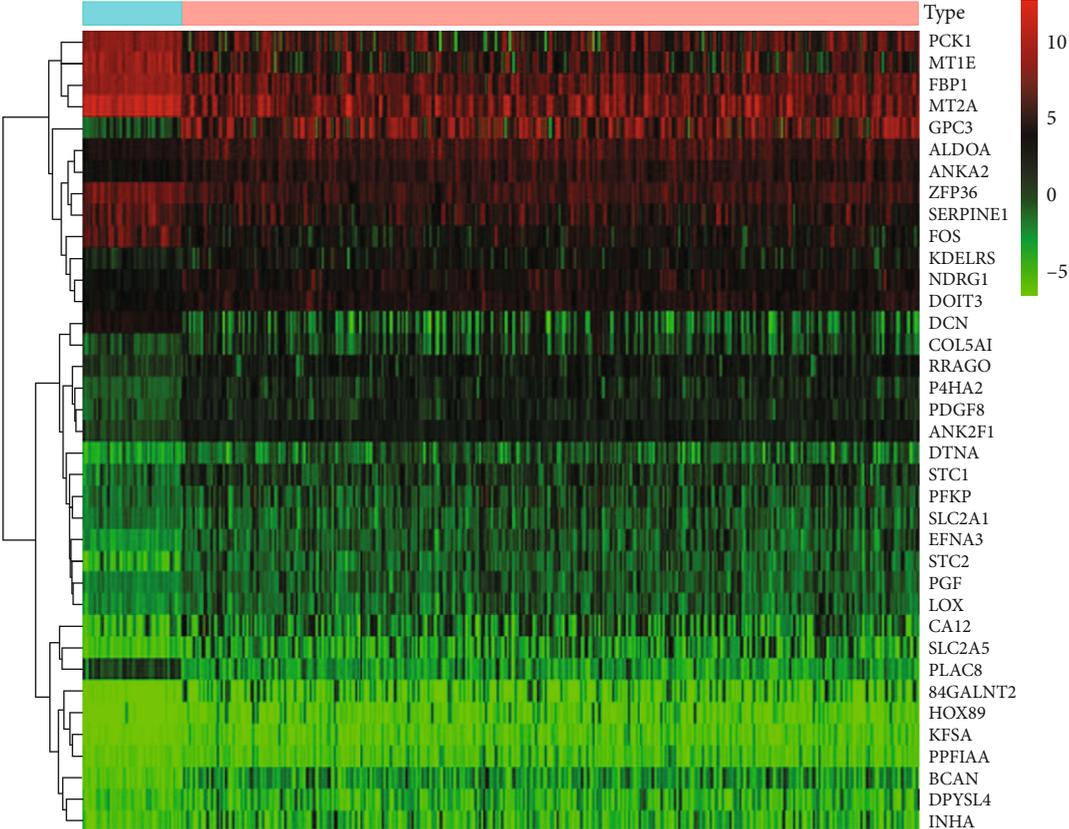
To further verify the hypoxia-related gene model and whether the PP could be used as an independent predictor of prognosis for HCC patients, we divided the TCGA data set into a training set and a validation set, and performed survival analysis on the hypoxia-related signature in the two data sets to illustrate the robustness of the PP. Besides, we introduced several following indicators closely related to the clinic into the model as covariables for multivariate Cox regression analysis: PP and age were coded as continuous variables, gender (male or female), tumor grade (high or low), AJCC TNM stage (I = 1, II = 2, III = 3, and IV = 4), tumor stage [1–4], lymph node metastasis (positive or negative), and distant metastasis (positive or negative).

2.4. Identification of Potential Therapeutic Agents. Connectivity map (CMap) was used to identify potential therapeutic agents for HCC. It is a program that predicts potential drugs and produces biological states by encoding specific gene expression characteristics [14]. 28 upregulated and 9 down-regulated HRGs were used to check the CMap database. Finally, the enrichment score representing similarity was calculated, ranging from -1 to 1. The positive connectivity score indicates that drugs can induce the biological phenomena queried in human cell lines. In contrast, a negative connectivity score indicates that the drug reverses the requested biological characteristics and has potential therapeutic value. With $p < 0.05$ as the cut-off value, the connectivity scores of different instances were filtered out. We carried out studies on these potentially effective drugs in PubChem.

2.5. Gene Set Cancer Analysis (GSCA). To further explore the underlying molecular mechanisms involved in these modeling genes, GSCA (<http://bioinfo.life.hust.edu.cn/web/GSCALite/>) [15] was used to obtain the activity of 10 cancer-related pathways. Genes were divided into high and low group by median expression. Student's T test was used to define the difference of pathway activity score (PAS) between groups. p value was adjusted by FDR (FDR ≤ 0.05 is considered as significant).



(a)



Type
N
T

(b)

FIGURE 1: Continued.

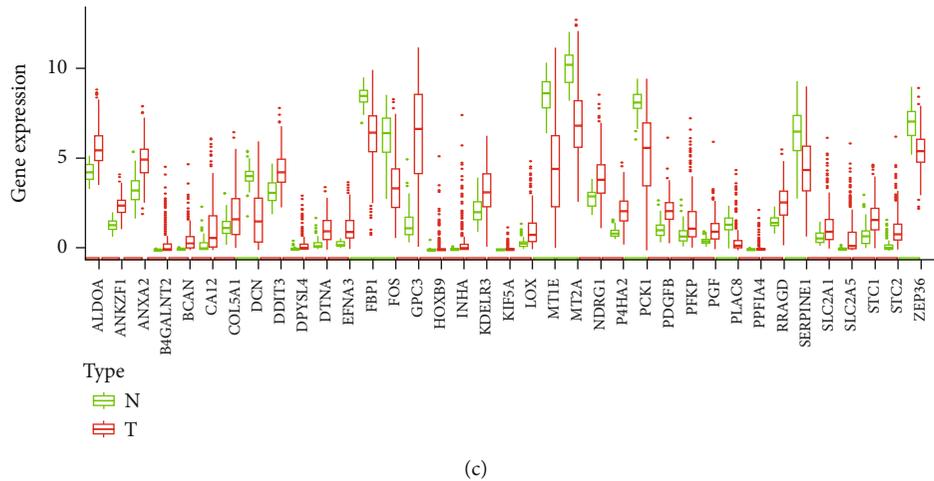
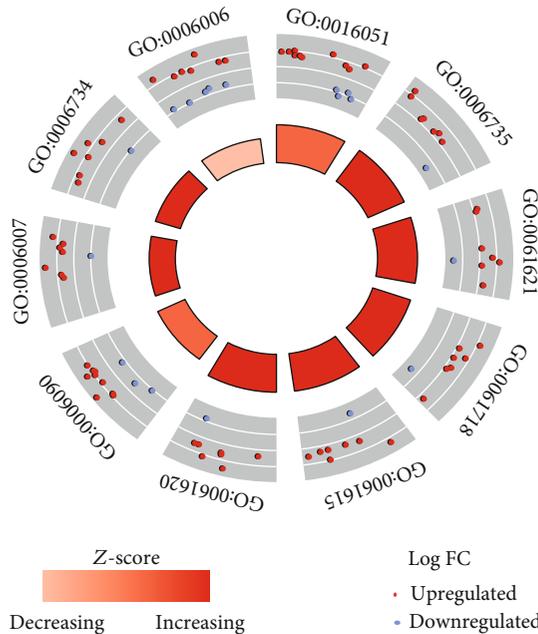


FIGURE 1: Differentially expressed HRGs between HCC and para-carcinoma tissues. (a) The volcano plot for the 200 HRGs from the TCGA data portal. 37 significantly different genes were labeled in the figure. Red indicates high expression and green low expression. (b) Hierarchical clustering of differentially expressed HRG expression levels. (c) The expression patterns of box plots displayed 37 differentially expressed HRGs in HCC types and paired nontumor samples. Each red dot represents a distinct tumor sample and green represents a nontumor sample.



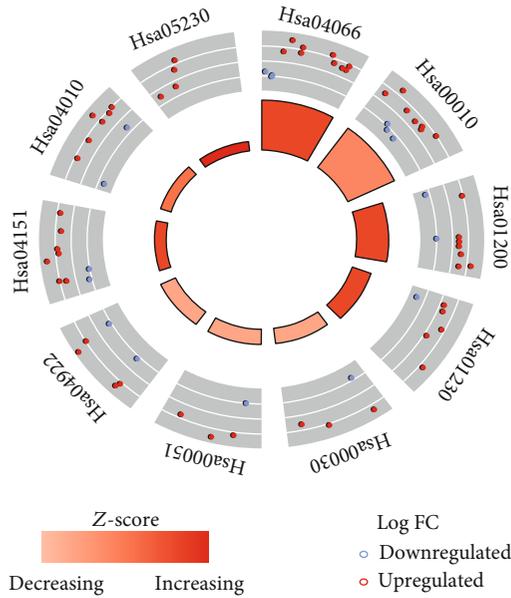
ID	Description
GO:0016051	Carbohydrate biosynthetic process
GO:0006735	NADH regeneration
GO:0061621	Canonical glycolysis
GO:0061718	Glucose catabolic process to pyruvate
GO:0061615	Glycolytic process through fructose-6-phosphate
GO:0061620	Glycolytic process through glucose-6-phosphate
GO:0006090	Pyruvate metabolic process
GO:0006007	Glucose catabolic process
GO:0006734	NADH metabolic process
GO:0006006	Glucose metabolic process

FIGURE 2: The circle plot of the top 10 GO terms function enrichment of differentially expressed HRGs.

2.6. *Statistical Analysis.* All statistical analyses were conducted using SPSS 22.0 (Chicago, IL, USA) and R 3.6.0 (<https://www.r-project.org/>); Adobe Illustrator CS 5 (San Jose, California, USA) was performed to draw plots. Univariate Cox regression analysis was used for illustrating the correlation between genes expression profiles and OS. The multivariate Cox proportional hazard regression model was used for performing multivariate analysis and construct prognostic models. Receiver-operating characteristic (ROC) curve and the corresponding area under the ROC curve (AUC) were applied for verifying that PP was accurate. All tests were performed bilaterally, $p < 0.05$ indicates a difference with statistical significance.

3. Results

3.1. *Differentially Expressed HRGs.* A total of 374 HCC tissue samples and 50 para-carcinoma tissue specimens were obtained for RNA-seq and clinical data from TCGA. A list of 200 HRGs involved in hypoxic regulatory pathways reported in the literature was extracted from the Hallmark hypoxic gene sets in the MSigDB database. According to the limitation of a $FDR < 0.05$ and $|\log_2(\text{Fold Change})| > 1.5$, our study screened 37 genes with significant differential expression, of which 28 were upregulated and 9 were downregulated (Figures 1(a) and 1(b)). Then, we visualized the expression patterns of these HRGs between HCC and normal



ID	Description
Hsa04066	HIF-1 signaling pathway
Hsa00010	Glycolysis/Gluconeogenesis
Hsa01200	Carbon metabolism
Hsa01230	Biosynthesis of amino acids
Hsa00030	Pentose phosphate pathway
Hsa00051	Fructose and mannose metabolism
Hsa04922	Glucagon signaling pathway
Hsa04151	PI3K-Akt signaling pathway
Hsa04010	MAPK signaling pathway
Hsa05230	Central carbon metabolism in cancer

FIGURE 3: The circle plot of the top 10 enriched KEGG pathways of differentially expressed HRGs.

tissues by making boxplot (Figure 1(c)). The boxplot showed expression patterns that 28 upregulated genes (*ALDOA*, *ANKZF1*, *ANXA2*, *B4GALNT2*, *BCAN*, *CA12*, *COL5A1*, *DDIT3*, *DPYSL4*, *DTNA*, *EFNA3*, *GPC3*, *HOXB9*, *INHA*, *KDEL3*, *KIF5A*, *LOX*, *NDRG1*, *P4HA2*, *PDGFB*, *PFKP*, *PGF*, *PPFIA4*, *RRAGD*, *SLC2A1*, *SLC2A5*, *STC1*, and *STC2*) and 9 downregulated genes (*DCN*, *FBP1*, *FOS*, *MT1E*, *MT2A*, *PCK1*, *PLAC8*, *SERPINE1*, and *ZFP36*).

3.2. Functional Annotation and Enrichment Analysis of Differentially Expressed HRGs. To further understand the biologically relevant information of these 37 differential HRGs, we performed functional enrichment and enrichment analysis. The GO term function (Figure 2) and KEGG pathway enrichment (Figure 3) of these genes were reviewed, respectively. We found the results that the GO terms for biological processes were carbohydrate biosynthetic process, NADH regeneration, canonical glycolysis, glucose catabolic process to pyruvate, and so on. The detailed results of all the above gene enrichment are shown in Figure 2.

Through the enrichment analysis function in the KEGG pathway, we found many important pathways associated with these genes such as HIF-1 signaling pathway, glycolysis/gluconeogenesis, carbon metabolism, MAPK signaling pathway, and PI3K-Akt signaling pathway (Figure 3). As shown in the circle plot, the change from the Z-score indicated mostly related signaling pathways were more inclined to be increased.

3.3. Identification of Prognostic HRGs. In order to build a model related to the prognosis of HCC, we selected patients with follow-up period of more than 1 month from this data set. Finally, we obtained 349 patients for the analysis and establishment of the prognostic predictor (PP). Prognostic HRGs were computed with univariate Cox regression model by SPSS 22.0. The relationship between the expression

TABLE 1: The associated HRGs with OS for univariate Cox regression model in HCC patients.

ID	HR	HR.95% CI	p value
HDLBP	1.76	1.19-2.60	<0.01
SAP30	1.92	1.47-2.51	<0.01
NDRG1	1.34	1.17-1.54	<0.01
ADORA2B	1.56	1.16-2.09	<0.01
PFKP	1.26	1.13-1.42	<0.01
DPYSL4	1.55	1.12-2.14	<0.01
PFKFB3	1.21	1.07-1.37	<0.01
SLC2A1	1.62	1.38-1.90	<0.01
VEGFA	1.43	1.14-1.80	<0.01
HMOX1	1.26	1.09-1.45	<0.01
SDC3	1.37	1.11-1.70	<0.01
ATP7A	2.01	1.28-3.16	<0.01
PGK1	1.45	1.17-1.79	<0.01
ERO1A	1.48	1.17-1.87	<0.01
XPNPEP1	2.03	1.32-3.12	<0.01
LDHA	2.09	1.60-2.73	<0.01
ENO2	1.29	1.10-1.52	<0.01
GRHPR	0.73	0.60-0.91	<0.01
UGP2	0.74	0.60-0.92	<0.01
SLC6A6	1.29	1.08-1.53	<0.01
MAP3K1	1.46	1.11-1.92	<0.01
BNIP3L	1.43	1.12-1.82	<0.01
TPI1	1.61	1.24-2.08	<0.01
RRAGD	1.32	1.07-1.62	<0.01
TES	1.28	1.07-1.54	<0.01
SERPINE1	1.16	1.05-1.28	<0.01
JMJD6	1.77	1.32-2.37	<0.01

HR: hazard ratio; CI: confidence interval.

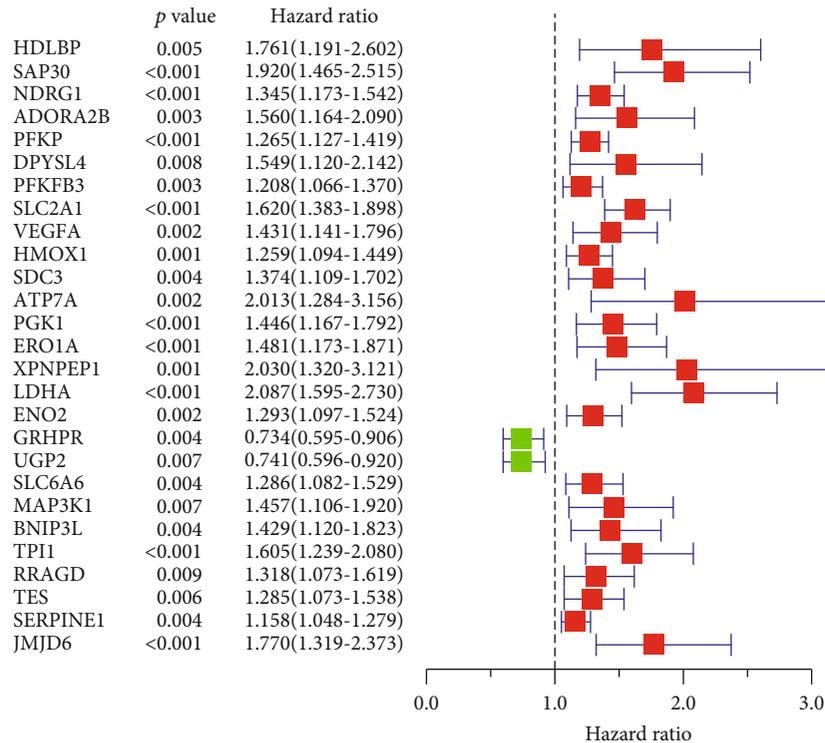


FIGURE 4: The relationships between the expression profiles of related genes (HRGs) and OS (overall survival). In these 27 HRGs, red squares display risk factors, and the other 2 green squares display protective factors.

profiles of HRGs and OS was evaluated using TCGA data and the Hallmark hypoxic gene sets in the MSigDB database, resulting in 27 prognosis-related HRGs ($HR > 1$ or $HR < 1$, $p < 0.01$) in Table 1.

By mapping the forest of HRGs and OS (Figure 4), we can intuitively understand the role of these genes in cancer prognosis. Of these, 25 genes (*HDLBP*, *SAP30*, *NDRG1*, *ADORA2B*, *PFKP*, *DPYSL4*, *PFKFB3*, *SLC2A1*, *VEGFA*, *HMOX1*, *SDC3*, *ATP7A*, *PGK1*, *ERO1A*, *XPNPEP1*, *LDHA*, *ENO2*, *SLC6A6*, *MAP3K1*, *BNIP3L*, *TPI1*, *RRAGD*, *TES*, *SERPINE1*, and *JMJD6*) were risk factors for cancer, and 2 genes (*GRHPR* and *UGP2*) were protective factors.

In order to better evaluate the prognosis and survival time of patients, an optimal equation for the prognostic predictor (PP) was further conducted by multivariate Cox proportional hazard regression model. Finally, 12 genes (*HDLBP*, *SAP30*, *PFKP*, *DPYSL4*, *SLC2A1*, *HMOX1*, *PGK1*, *ERO1A*, *LDHA*, *ENO2*, *SLC6A6*, and *TPI1*) were obtained and incorporated into the final model.

Among these genes, 9 genes (*HDLBP*, *SAP30*, *PFKP*, *DPYSL4*, *SLC2A1*, *HMOX1*, *ERO1A*, *LDHA*, and *TPI1*) were risk factors for HCC patients, and 3 genes (*PGK1*, *ENO2*, and *SLC6A6*) were protective factors (Table 2).

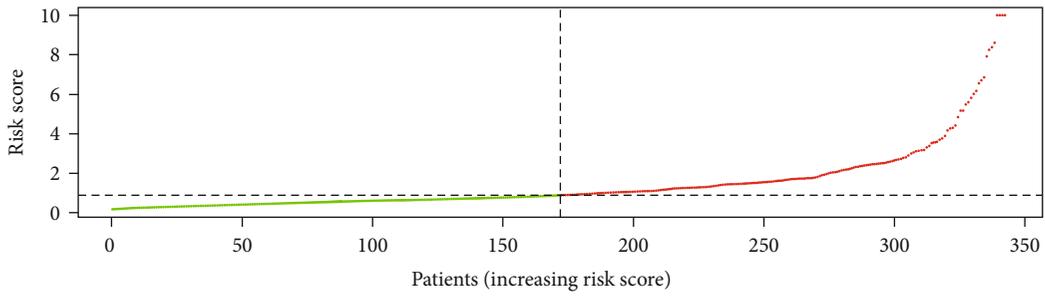
3.4. Construction and Definition of the PP. In the multivariate Cox proportional hazard regression model, Y (survival time and status) was obtained by multiplying the contribution weight of every gene and X (expression value of each gene) and then adding them. The formula of PP expressed as follows: $PP = (0.5003 \times \text{expression value of HDLBP}) + (0.4192$

TABLE 2: The associated HRGs with OS for multivariate Cox proportional hazard regression model.

ID	Coefficient	HR	HR.95% CI	<i>p</i> value
HDLBP	0.5003	1.6492	1.0947-2.4846	0.0167
SAP30	0.4192	1.5208	1.0984-2.1056	0.0116
PFKP	0.2345	1.2643	1.0443-1.5305	0.0162
DPYSL4	0.3718	1.4503	1.0289-2.0442	0.0338
SLC2A1	0.4166	1.5168	1.1191-2.0557	0.0072
HMOX1	0.1311	1.1401	0.9705-1.3394	0.1105
PGK1	-0.4244	0.6542	0.4637-0.9230	0.0157
ERO1A	0.3085	1.3614	0.9634-1.9239	0.0803
LDHA	0.5351	1.7077	1.2296-2.3716	0.0014
ENO2	-0.5150	0.5975	0.4269-0.8362	0.0027
SLC6A6	-0.2278	0.7963	0.6013-1.0547	0.1121
TPI1	0.3658	1.4417	1.0196-2.0387	0.0385

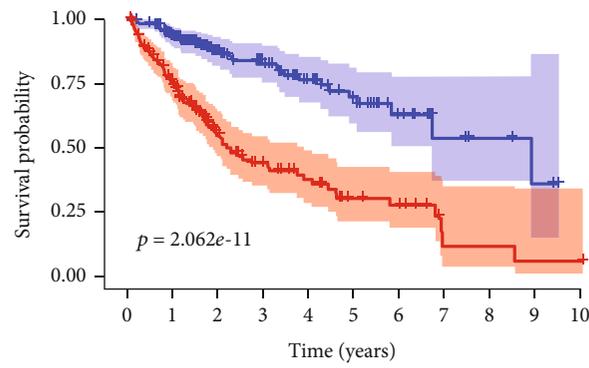
HR: hazard ratio; CI: confidence interval.

$\times \text{expression value of SAP30}) + (0.2345 \times \text{expression value of PFKP}) + (0.3718 \times \text{expression value of DPYSL4}) + (0.4166 \times \text{expression value of SLC2A1}) + (0.1311 \times \text{expression value of HMOX1}) + (-0.4244 \times \text{expression value of PGK1}) + (0.3085 \times \text{expression value of ERO1A}) + (0.5351 \times \text{expression value of LDHA}) + (-0.5150 \times \text{expression value of ENO2}) + (-0.2278 \times \text{expression value of SLC6A6}) + (0.3658 \times \text{expression value of TPI1})$. It was worth noting that the coefficients of the genes were positive; the expression of these genes may shorten the



- High risk
- Low risk

(a)

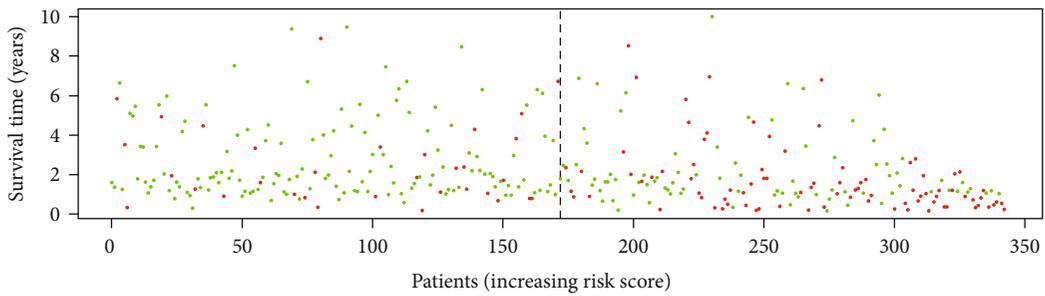


Risk	High risk	171	110	49	31	22	13	11	2	2	1	1
	Low risk	172	144	77	58	41	27	14	6	4	2	0
		0	1	2	3	4	5	6	7	8	9	10

Risk

- High risk
- Low risk

(b)



- Dead
- Alive

(c)

FIGURE 5: Continued.

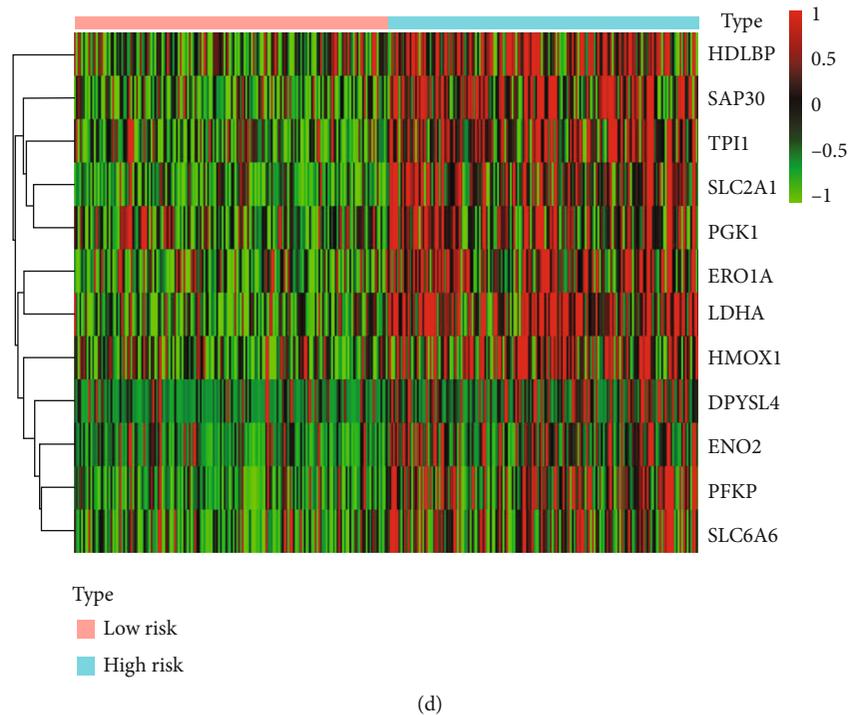


FIGURE 5: Hypoxia-related prognostic predictor (PP) of hepatocellular carcinoma patients. (a) High- and low-risk group with the median risk score as the cut-off point according to the risk score formula. Red indicates high risk and green low risk. (b) Kaplan-Meier plot represents that patients in the high-risk group had significantly shorter overall survival time than those in the low-risk group. (c) The overall survival of patients in the high- and low-risk group. The dotted line was the dividing line between high- and low-risk group of HCC patients. The right side of the dotted line represented the high-risk group, the left side represented the low-risk group, the red dots represented patients who had died, and the green dots represented patients who had still survived (d) The heat map shows the distribution of these 12 key gene expression profiles in the TCGA dataset.

OS of patients with HCC. Conversely, the negative coefficient of the gene means that these genes were beneficial for increasing the OS of HCC patients.

3.5. Validation the Value of PP in Evaluating Patients' Clinical Prognosis. To validate the performance of PP in evaluating patients' clinical prognosis, we divided patients into the high- and low-risk group with the median risk score as the cut-off point according to the risk score formula (Figure 5(a)). The K-M plots were plotted to analyze the different survival time between the two groups. The K-M analysis showed that patients in the high-risk group suffered significantly worse survival than those in the low-risk group (Figure 5(b)).

All these results indicated that the prediction ability of the equation was very well. Furthermore, the distribution of patients according to their prognostic risk was also showed (Figures 5(c) and 5(d)). We could see intuitively that the high-risk group had a much higher number of deaths than the low-risk group.

3.6. The Accuracy of PP Was Verified by Combining with Clinical Parameters. To further validate the accuracy of PP in predicting OS in HCC patients, our research performed the 1-, 3-, and 5-year ROC curves by integrating age, gender, tumor grade, and AJCC TNM stage. We got the AUC value of the 1-, 3-, and 5-year risk score were all >0.7 , which were

much bigger than other clinical parameters (Figure 6(a)). Thus, compared with other indicators, PP had important reference value in predicting patients' prognosis.

Furthermore, we used four-fifths of the samples that met the inclusion criteria from the TCGA data set as the training set, and the remaining one-fifth samples as the validation set. The results of the two data sets with survival analysis showed that the survival probability of the high-risk group was significantly lower than that of the low-risk group, and the difference was significant (Figures 6(b) and 6(c)). Risk score remained as an independent prognostic indicator for HCC patient in univariate and multivariate analyses, after adjusting for clinicopathological features such as age, gender, tumor grade, and AJCC TNM stage (HR = 1.484, 95%CI = 1.342 – 1.642, $p < 0.001$, and HR = 1.414, 95%CI = 1.258 – 1.588, $p < 0.001$, Figures 6(d) and 6(e)).

3.7. Relationship between PP and Clinical Parameters. To further clarify the important value of PP in clinical application, we explored the relationship between risk score and clinical indicators by using the "Beeswarm" software package. The relationship between risk score of the high- and low-risk group of HCC patients and the six clinical indicators (age, gender, tumor grade, and AJCC TNM stage) was verified by the "Beeswarm" software package; $p < 0.05$ was considered to be statistically significant. The final results showed that the risk score calculated by the PP model was consistent with

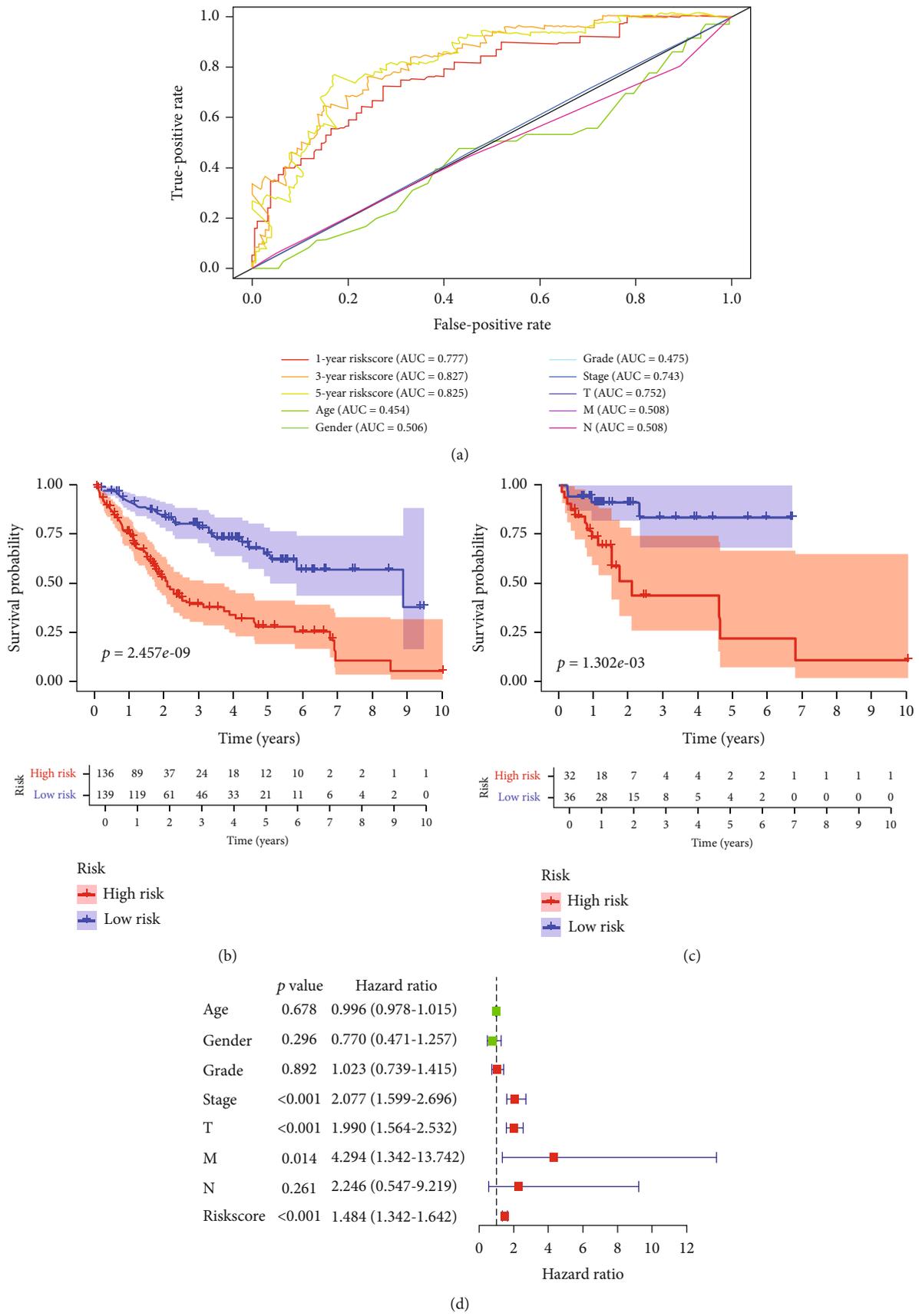


FIGURE 6: Continued.

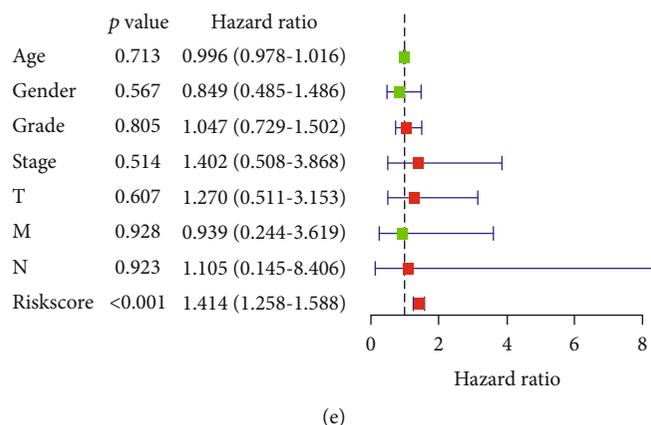


FIGURE 6: The accuracy of hypoxia-related prognostic predictor (PP) in hepatocellular carcinoma patients. (a) The 1-, 3-, and 5-year risk score ROC (receiver-operating characteristic) curve analysis by integrating age, gender, tumor grade, and AJCC TNM stage. Kaplan-Meier (K-M) method was used to plot the survival curve for the training set (b) and the validation set (c). Risk score with clinicopathological features in univariate analyses (d) and in multivariate analyses (e). Red squares displayed risk factors, and green squares displayed protective factors.

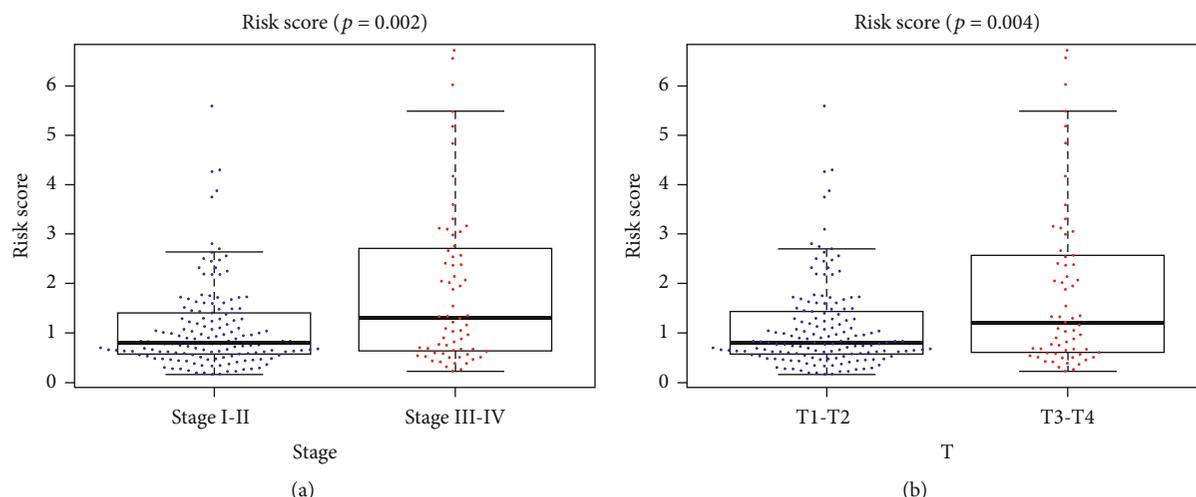


FIGURE 7: The relationship between risk score and clinical indicators. (a) The boxplot showed the relationship between risk score and tumor stage. (b) The boxplot showed the relationship between risk score and T stage.

the change in tumor stage (Figure 7(a)) and T stage (Figure 7(b)). The higher the patient's risk score, the higher the corresponding tumor stage and T stage, and the patient's prognosis was poor.

3.8. Potential Therapeutic Drug Screening. In order to screen out potential therapeutic drugs, we analyzed 28 upregulated and 9 downregulated HRGs with CMap. We got the top 4 of small molecules: deferoxamine, crotamiton, and talampicillin showed a highly positive correlation with HCC; lycorine showed a highly negative correlation with HCC. They all might become effective drugs for treating HCC. The 2D chemical structural formulas of the top 4 most promising molecular drugs from PubChem were shown in Figures 8(a)–8(d).

3.9. Gene Set Cancer Analysis (GSCA). In order to further explore the underlying molecular mechanisms of the 12 core genes involved in the PP, GSCA was used to obtain the activity of 10 cancer-related pathways. Figure 9 showed the

overall activity of these 12 key genes in 10 cancer-related pathways.

4. Discussion

HCC is a major contributor of death caused by cancer. Previously, many progresses have been made in understanding the epidemiology, risk factors, and molecular profiles of HCC [16]. However, incidence and HCC-specific mortality still continue to increase [17–19]. Although some progress has been made in molecular targeted therapy for HCC, the results are still unsatisfactory [20]. It is urgently needed to better understand the molecular mechanism which leads to the development of HCC and to search for targeted genes that play a key role in the prognosis of HCC for early intervention.

Exploration of hypoxia mechanism opens new perspectives for HCC [21]. In particular, the announcement of the 2019 Nobel Prize in physiology or medicine has increased researchers' enthusiasm for exploring the mechanism of

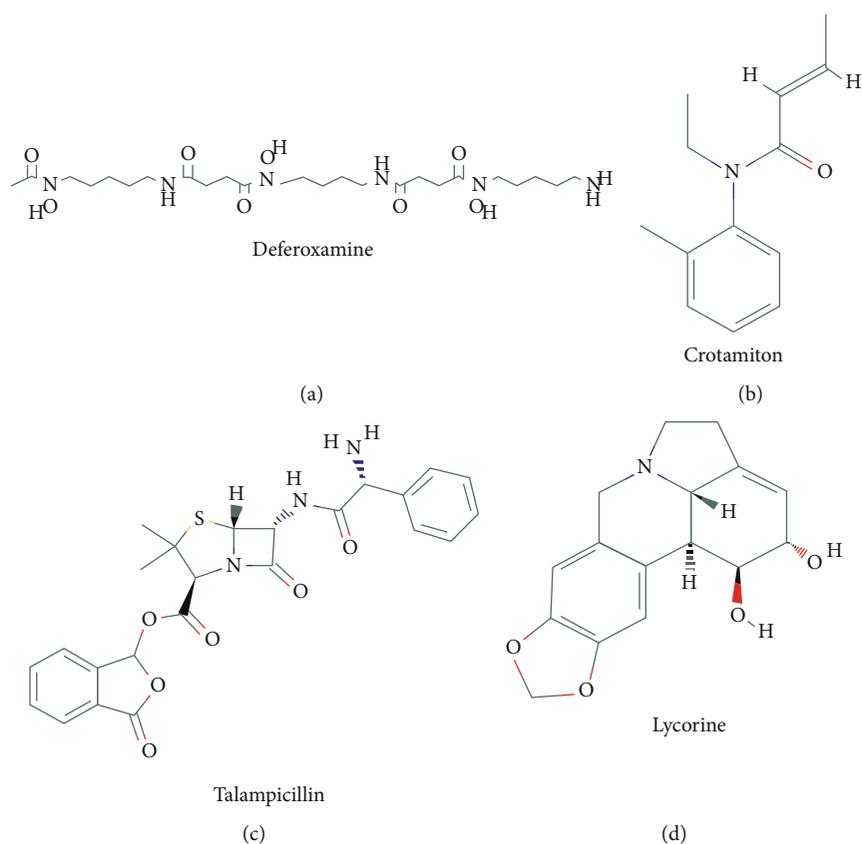


FIGURE 8: The structural formulas of the top 4 molecular drugs: (a) deferoxamine, (b) crotamiton, (c) talampicillin, and (d) lycorine.

hypoxia in the field of cancer. Increasing evidence indicates that hypoxia is closely related to the migration and malignant progression of HCC [8, 22]. However, most studies have only focused on hypoxia by studying signaling genes [23, 24]. With the development of bioinformatics and the continuous improvement of high-throughput sequencing technology, some large-scale databases, including TCGA and GEO, have emerged. These databases provide effective means for sorting out gene signatures. In the current study, we deeply mined all the differential genes which are significantly expressed in tumors and normal tissues. By integrating them with a large sample of clinical data, we aimed to select molecular biomarkers for detecting the prognosis of HCC patients.

We first screened 37 differentially expressed HRGs between HCC and para-carcinoma tissues. To discover the role of genes in HCC progression, GO and KEGG analysis of the differential expression genes was performed. Gene functional enrichment analysis suggested that these genes were mainly involved in carbohydrate biosynthetic process and HIF-1 signaling pathway. The process of carbohydrate biosynthesis was mainly related to the energy metabolism of cancer cells. HIF-1 and its related genes actively promoted HCC growth, HCC cell proliferation, and aggressive behavior, which was positively associated with the presence of intrahepatic metastasis and the histological grade of HCC [8, 25]. Our analysis confirmed that HIF-1 signaling pathway was the most critical process for hypoxia leading to poor prognosis of HCC, which is consistent with the findings in

most literatures [26, 27]. As such, these genes could also act as clinical biomarkers for monitoring metastasis, assessing survival, and potential drug targets. Bioinformatics analysis provided several clues intervening the occurrence and development of HCC via several hypoxic pathways. The univariate Cox regression analysis revealed that 27 HRGs were associated with OS in the TCGA database. Further multivariate Cox regression analysis helped us determine 12 vital prognostic HRGs (*HDLBP*, *SAP30*, *PFKP*, *DPYSL4*, *SLC2A1*, *HMOX1*, *PGK1*, *ERO1A*, *LDHA*, *ENO2*, *SLC6A6*, and *TPI1*) to construct of the PP. There are already several existed reports about partial of these genes.

HDLBP (high-density lipoprotein-binding protein), a positively regulated gene, is closely related to the cancer process [28]. HDLBP can promote HCC cell proliferation and tumorigenesis. Consistent with many other studies, HDLBP was overexpressed in other types of cancers [29]. Intriguingly, in breast cancer, HDLBP may be a tumor suppressor to accelerate the degradation and inhibit the translation of the *c-fms* protooncogene mRNA [30]. These observations suggested that the mechanism of HDLBP (either inhibit or promote carcinogenesis) was still unclear. SAP30 (serum amyloid P30) was a very sensitive indicator of liver disease [31]. It has been demonstrated that SAP upregulated in serum samples from HCC patients. HMOX1 (heme oxygenase 1) is an isoenzyme of heme oxygenase (HO). HO is a microsomal enzyme system that degrades heme. Their main biological role is to degrade heme to produce carbon

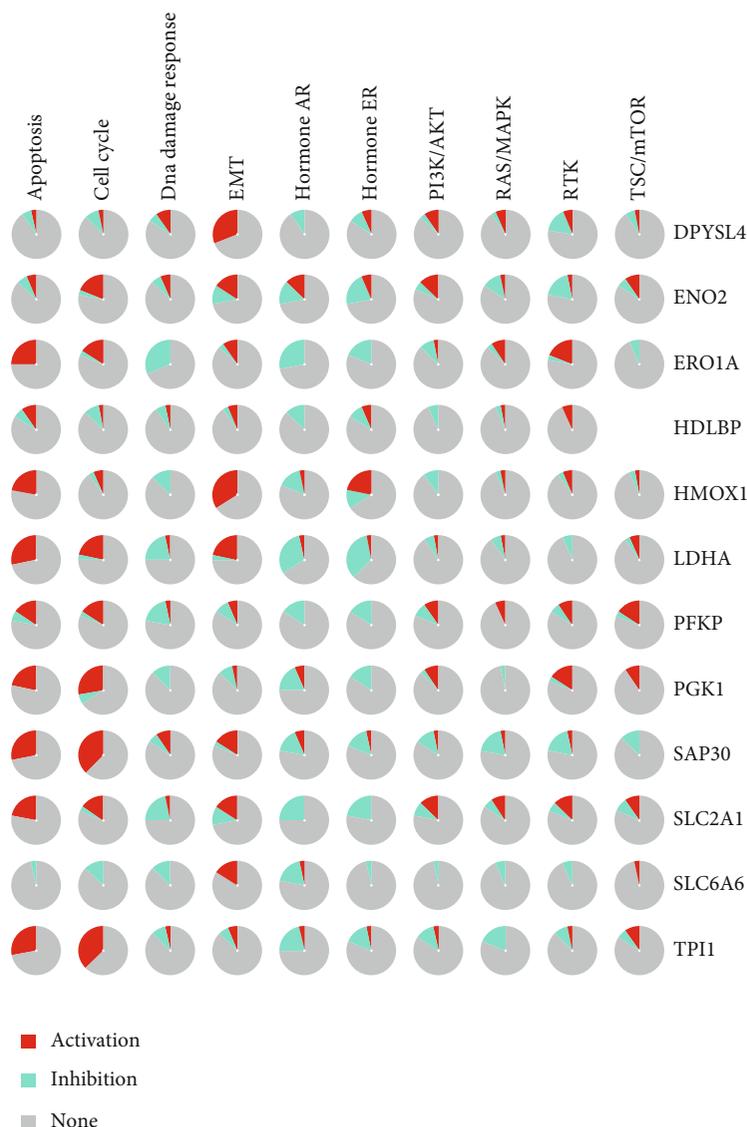


FIGURE 9: The activity of 10 cancer-related pathways with GSCA.

monoxide (CO) and biliverdin and release iron ions. HMOX1 has been proved to be an oncogene and plays an important role in the development of various cancers. Inhibition of HMOX1 degradation can definitely promote HCC proliferation. LDHA (lactate dehydrogenase A) is a vital enzyme responsible for cancer growth and energy metabolism via the aerobic glycolytic pathway. Inhibition of LDHA could inhibit tumor-initiating cell survival and proliferation, which indicated that LDHA may be a potential therapeutics target [32]. PGK1 (phosphoglycerate kinase 1) is an important enzyme in the metabolic glycolysis pathway. Many articles have observed a significant overexpression of PGK1 in HCC tissues and a negative correlation between PGK1 expression and HCC patient survival. Depletion of PGK1 dramatically reduced cancer cell proliferation and tumorigenesis [33, 34]. These findings demonstrated a novel role of PGK1 in HCC progression. *ENO2* (Enolase 2) encodes an enolase isoenzyme which is regarded as a sensitive biomarker for neuroendocrine tumors [35]. Additionally,

ENO2 was believed to be an indicator in the diagnosis and prognosis rather than a potential target for therapy in renal tumor patients [36]. In our study, *ENO2* was defined as a negative regulatory gene. It provided new therapeutic targets and biomarkers for HCC prognosis through inhibiting its expression to improve the prognosis of HCC. Other genes were rarely reported; their roles and underlying mechanism in cancer still need to be further explored.

According to the current study, some features of HCC prognosis based on expression profile have been found by large public databases. However, the purpose of these studies was often limited to mining new molecular markers but neglecting routine clinical parameters [37, 38]. Thus, we integrated patient clinical information to establish a PP model for assessing prognosis. Firstly, we verified the accuracy of the model through the K-M survival analysis. The risk score was calculated by PP formula, and the median was taken as the cut-off point dividing the high- and low-risk group, and the survival difference was significant (Figure 5(a)). Secondly,

we performed the 1-, 3-, and 5-year risk score ROC curve analysis by integrating age, gender, tumor grade, and AJCC TNM stage and found that the AUC value of 1-, 3-, and 5-year risk score were all much bigger than other clinical parameters (Figure 6(a)). Besides, we divided the TCGA data set into a training set and a validation set and performed survival analysis on the hypoxia-related signature in the two data sets to illustrate the robustness of the PP. The results of the two data sets with survival analysis showed that the survival probability of the high-risk group was significantly lower than that of the low-risk group, and the difference was significant (Figures 6(b) and 6(c)). Similarly, we found that the risk score was still an independent indicator for evaluating the prognosis of patients with HCC through the process of univariate and multivariate analysis (Figures 6(d) and 6(e)). At last, by comprehensively analyzing the risk score and the patient's tumor TNM stage, we found that the greater the risk score calculated by PP, the higher the patient's tumor grade and the worse the prognosis. All of the above fully demonstrated the accuracy of our constructed model of PP and its important value in evaluating patients' prognosis.

The results of CMap demonstrated the therapeutic agents closely associated with HCC. Substantial evidence suggests that iron overload enhances cancer growth and metastasis. As a commonly used iron chelator, high-dose deferoxamine therapy could inhibit tumor migration by destroying iron homeostasis, arresting the cell cycle and enhancing apoptosis of hepatocellular carcinoma and hepatoblastoma cell lines [39]. The iron chelator deferoxamine has been proved that it could prevent liver injury as well as the development of preneoplastic lesions in rats. In particular, as a potential anticancer drug, it has been used in clinical studies to confirm its efficacy in patients with advanced hepatocellular carcinoma [40]. Lycorine is a multifunctional bioactive compound and powerful anticancer agent against various cancer cell lines. Focusing on hepatocellular carcinoma, lycorine can significantly induce apoptosis and autophagy of hepatocellular carcinoma cells both in vivo and in vitro [41]. The possible molecular mechanism involved the TCRP1/Akt/mTOR pathway. Lycorine inhibited TCRP1 to reduce the phosphorylation level of Akt and inhibited Akt/mTOR signal transduction and finally mediated the apoptosis and autophagy of xenograft hepatocellular tumors without remarkable toxicity. Similarly, research by Hu et al. suggested that lycorine could be used to treat colorectal cancer by inducing autophagy-associated apoptosis [42]. Additionally, lycorine could inhibit the proliferation of SGC-7901 gastric cancer cells and promote cell apoptosis [43]. The specific mechanism of treatment was exactly consistent with the results of GSCA for modeling gene pathway analysis. However, there were few reports about crotamiton and talampicillin in the treatment of cancer. Their anticancer mechanisms need to be further explored and studied.

In conclusion, through a comprehensive bioinformatics analysis with HRG expression profiles, our study firstly reported 12 prognostic HRGs and constructed the model of the PP. This comprehensive research of multiple databases helps us gain insight into the biological properties of HCC

and provides deferoxamine, crotamiton, talampicillin, and lycorine as potential molecular drugs to fight against HCC.

Data Availability

The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request.

Conflicts of Interest

The authors have no conflicts of interest to declare.

Authors' Contributions

HG performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft. JZ performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft. YZ performed the experiments, analyzed the data, and approved the final draft. ZW analyzed the data, prepared figures and/or tables, and approved the final draft. LL performed the experiments, analyzed the data, and approved the final draft. WZ performed the experiments, prepared figures and/or tables, and approved the final draft. All authors read and approved the final manuscript.

Acknowledgments

We are grateful to the patient and his family for their contributions to the study. We would like to thank the TCGA, MsigDB, CMap, and GSCALite databases for the availability of the data. This study was supported by the National Natural Science Foundation of China (No. 81904183), the Project of Health Commission of Shanxi Province (No. 2018078), the Natural Science Foundation of Shanxi Province (No. 201801D121300), and the cultivation project of Science and Technology Innovation Ability of Shanxi University of Chinese Medicine (2019PY-017).

References

- [1] P. Vaupel and A. Mayer, "Hypoxia in cancer: significance and impact on clinical outcome," *Cancer Metastasis Reviews*, vol. 26, no. 2, pp. 225–239, 2007.
- [2] X. X. Xiong, X. Y. Qiu, D. X. Hu, and X. Q. Chen, "Advances in hypoxia-mediated mechanisms in hepatocellular carcinoma," *Molecular Pharmacology*, vol. 92, no. 3, pp. 246–255, 2017.
- [3] S. Osinsky, M. Zavelevych, and P. Vaupel, "Tumor hypoxia and malignant progression," *Experimental Oncology*, vol. 31, no. 2, pp. 80–86, 2009.
- [4] P. Vaupel and A. Mayer, "Hypoxia and anemia: effects on tumor biology and treatment resistance," *Transfusion Clinique et Biologique*, vol. 12, no. 1, pp. 5–10, 2005.
- [5] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.

- [6] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2018," *CA: a Cancer Journal for Clinicians*, vol. 68, no. 1, pp. 7–30, 2018.
- [7] H. Pan, X. Fu, and W. Huang, "Molecular mechanism of liver cancer," *Anti-Cancer Agents in Medicinal Chemistry*, vol. 11, no. 6, pp. 493–499, 2011.
- [8] Z. Song, T. Liu, J. Chen et al., "HIF-1 α -induced RIT1 promotes liver cancer growth and metastasis and its deficiency increases sensitivity to sorafenib," *Cancer Letters*, vol. 460, pp. 96–107, 2019.
- [9] Y. H. Su, A. K. Kim, and S. Jain, "Liquid biopsies for hepatocellular carcinoma," *Translational Research*, vol. 201, pp. 84–97, 2018.
- [10] Cancer Genome Atlas Research Network, "Electronic address wbe, Cancer Genome Atlas Research N. Comprehensive and integrative genomic characterization of hepatocellular carcinoma," *Cell*, vol. 169, pp. 1327–1341.e23, 2017.
- [11] A. Liberzon, C. Birger, H. Thorvaldsdottir, M. Ghandi, J. P. Mesirov, and P. Tamayo, "The Molecular Signatures Database (MSigDB) hallmark gene set collection," *Cell Systems*, vol. 1, no. 6, pp. 417–425, 2015.
- [12] A. Subramanian, P. Tamayo, V. K. Mootha et al., "Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 43, pp. 15545–15550, 2005.
- [13] G. Yu, L. G. Wang, Y. Han, and Q. Y. He, "clusterProfiler: an R package for comparing biological themes among gene clusters," *OMICS*, vol. 16, no. 5, pp. 284–287, 2012.
- [14] J. Lamb, E. D. Crawford, D. Peck et al., "The Connectivity Map: using gene-expression signatures to connect small molecules, genes, and disease," *Science*, vol. 313, no. 5795, pp. 1929–1935, 2006.
- [15] C. J. Liu, F. F. Hu, M. X. Xia, L. Han, Q. Zhang, and A. Y. Guo, "GSCALite: a web server for gene set cancer analysis," *Bioinformatics*, vol. 34, no. 21, pp. 3771–3772, 2018.
- [16] J. D. Yang, P. Hainaut, G. J. Gores, A. Amadou, A. Plymoth, and L. R. Roberts, "A global view of hepatocellular carcinoma: trends, risk, prevention and management," *Nature Reviews Gastroenterology & Hepatology*, vol. 16, no. 10, pp. 589–604, 2019.
- [17] J. Zhu, T. Yin, Y. Xu, and X. J. Lu, "Therapeutics for advanced hepatocellular carcinoma: recent advances, current dilemma, and future directions," *Journal of Cellular Physiology*, vol. 234, no. 8, pp. 12122–12132, 2019.
- [18] M. Yarchoan, P. Agarwal, A. Villanueva et al., "Recent developments and therapeutic strategies against hepatocellular carcinoma," *Cancer Research*, vol. 79, no. 17, pp. 4326–4330, 2019.
- [19] C. Fitzmaurice, C. Allen, R. M. Barber et al., "Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study," *JAMA Oncology*, vol. 3, pp. 524–548, 2017.
- [20] W. Huang, J. Liu, J. Yan et al., "LncRNA LINC00470 promotes proliferation through association with NF45/NF90 complex in hepatocellular carcinoma," *Human Cell*, vol. 33, no. 1, pp. 131–139, 2020.
- [21] C. Bowyer, A. L. Lewis, A. W. Lloyd, G. J. Phillips, and W. M. Macfarlane, "Hypoxia as a target for drug combination therapy of liver cancer," *Anti-Cancer Drugs*, vol. 28, no. 7, pp. 771–780, 2017.
- [22] L. Y. Guo, P. Zhu, and X. P. Jin, "Association between the expression of HIF-1 α and VEGF and prognostic implications in primary liver cancer," *Genetics and Molecular Research*, vol. 15, no. 2, 2016.
- [23] S. Carbajo-Pescador, R. Ordoñez, M. Benet et al., "Inhibition of VEGF expression through blockade of Hif1 α and STAT3 signalling mediates the anti-angiogenic effect of melatonin in HepG2 liver cancer cells," *British Journal of Cancer*, vol. 109, no. 1, pp. 83–91, 2013.
- [24] S. J. Myung, J. H. Yoon, and S. J. Yu, "STAT3 & cytochrome P450 2C9: a novel signaling pathway in liver cancer stem cells," *Bio-medicine & Pharmacotherapy*, vol. 66, no. 8, pp. 612–616, 2012.
- [25] Y. Lei, Q. Hu, and J. Gu, "Expressions of carbohydrate response element binding protein and glucose transporters in liver cancer and clinical significance," *Pathology & Oncology Research*, vol. 26, no. 2, pp. 1331–1340, 2020.
- [26] X. Ning, Y. Wang, W. Yan, G. Li, and N. Sang, "Chitin synthesis inhibitors promote liver cancer cell metastasis via interfering with hypoxia-inducible factor 1 α ," *Chemosphere*, vol. 206, pp. 231–237, 2018.
- [27] F. Hu, X. Deng, X. Yang et al., "Hypoxia upregulates Rab11-family interacting protein 4 through HIF-1 α to promote the metastasis of hepatocellular carcinoma," *Oncogene*, vol. 34, no. 49, pp. 6007–6017, 2015.
- [28] W. Zhou, L. Zhao, H. Yuan et al., "A new small cell lung cancer biomarker identified by cell-SELEX generated aptamers," *Experimental Cell Research*, vol. 382, no. 2, p. 111478, 2019.
- [29] W. L. Yang, L. Wei, W. Q. Huang et al., "Vigilin is overexpressed in hepatocellular carcinoma and is required for HCC cell proliferation and tumor growth," *Oncology Reports*, vol. 31, no. 5, pp. 2328–2334, 2014.
- [30] H. H. Woo, X. Yi, T. Lamb et al., "Posttranscriptional suppression of proto-oncogene c-fms expression by vigilin in breast cancer," *Molecular and Cellular Biology*, vol. 31, no. 1, pp. 215–225, 2010.
- [31] G. Ferrín, I. Ranchal, C. Llamozza et al., "Identification of candidate biomarkers for hepatocellular carcinoma in plasma of HCV-infected cirrhotic patients by 2-D DIGE," *Liver International*, vol. 34, no. 3, pp. 438–446, 2014.
- [32] T. W. Chung, E. Y. Kim, C. W. Han et al., "Machilin A inhibits tumor growth and macrophage M2 polarization through the reduction of lactic acid," *Cancers*, vol. 11, no. 7, p. 963, 2019.
- [33] H. Hu, W. Zhu, J. Qin et al., "Acetylation of PGK1 promotes liver cancer cell proliferation and tumorigenesis," *Hepatology*, vol. 65, no. 2, pp. 515–528, 2016.
- [34] H. Xie, G. Tong, Y. Zhang, S. Liang, K. Tang, and Q. Yang, "PGK1 drives hepatocellular carcinoma metastasis by enhancing metabolic process," *International Journal of Molecular Sciences*, vol. 18, no. 8, p. 1630, 2017.
- [35] M. A. Isgro, P. Bottoni, and R. Scatena, "Neuron-specific enolase as a biomarker: biochemical and clinical aspects," *Advances in Experimental Medicine and Biology*, vol. 867, pp. 125–143, 2015.
- [36] T. Luo, X. Chen, S. Zeng et al., "Bioinformatic identification of key genes and analysis of prognostic values in clear cell renal cell carcinoma," *Oncology Letters*, vol. 16, no. 2, pp. 1747–1757, 2018.
- [37] Z. Wan, X. Zhang, Y. Luo, and B. Zhao, "Identification of hepatocellular carcinoma-related potential genes and pathways

- through bioinformatic-based analyses," *Genetic Testing and Molecular Biomarkers*, vol. 23, no. 11, pp. 766–777, 2019.
- [38] Y. Zheng, K. Yu, C. Huang et al., "Integrated bioinformatics analysis reveals role of the LINC01093/miR-96-5p/ZFAND5/NF- κ B signaling axis in hepatocellular carcinoma," *Experimental and Therapeutic Medicine*, vol. 18, no. 5, pp. 3853–3860, 2019.
- [39] E. Tabor and C. M. Kin, "Inhibition of human hepatocellular carcinoma and hepatoblastoma cell lines by deferoxamine," *Journal of Medical Virology*, vol. 34, no. 1, pp. 45–50, 1991.
- [40] T. Yamasaki, S. Terai, and I. Sakaida, "Deferoxamine for advanced hepatocellular carcinoma," *The New England Journal of Medicine*, vol. 365, no. 6, pp. 576–578, 2011.
- [41] H. Y. Yu, Y. L. Qiu, X. Pang et al., "Lycorine promotes autophagy and apoptosis via TCRP1/Akt/mTOR axis inactivation in human hepatocellular carcinoma," *Molecular Cancer Therapeutics*, vol. 16, no. 12, pp. 2711–2723, 2017.
- [42] M. Hu, Z. Yu, P. Mei et al., "Lycorine induces autophagy-associated apoptosis by targeting MEK2 and enhances vemurafenib activity in colorectal cancer," *Aging*, vol. 12, no. 1, pp. 138–155, 2020.
- [43] X. Chen, X. Long, Y. Liao, and J. Zheng, "Lycorine inhibits the proliferation and promotes apoptosis of SGC-7901 gastric cancer cells by downregulating expression of cyclin D1," *Xi Bao Yu Fen Zi Mian Yi Xue Za Zhi*, vol. 34, no. 4, pp. 354–358, 2018.