

Screening hypoxia-related genes as prognostic biomarkers and modeling the individualized prognostic predictor in hepatocellular carcinoma

Hao Guo

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

Jun Yan Zhang

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

Shu Jing Lv

Department of Anesthesiology, The First Central Clinical College, Tianjin Medical University, Tianjin 300192, China

Zhi Wang

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

Wei Wei Zhang

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

Kun Li Liu

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

Yan Cheng

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

Jing Zhou (✉ zhoujingsx07@163.com)

Shanxi Province Academy Of Traditional Chinese Medicine, Shanxi Province Hospital of Traditional Chinese Medicine

Research article

Keywords: hypoxia, prognostic predictor, hypoxic related genes, hepatocellular carcinoma, The Cancer Genome Atlas

Posted Date: December 23rd, 2019

DOI: <https://doi.org/10.21203/rs.2.19456/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Abstract

Background Hypoxia closely relates to malignant progression and appears to be prognostic for outcome in hepatocellular carcinoma (HCC). Our research aims to mine the Hypoxic related genes (HRGs) on the role of clinical prognosis in HCC. Moreover, we construct and define a model of prognostic predictor (PP) to estimate and improve prognosis of HCC patients.

Results 37 differentially expressed HRGs were obtained. It contained 28 up-regulated and 9 down-regulated 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 of the 10 key genes (HDLBP, SAP30, PFKP, DPYSL4, SLC2A1, PGK1, ERO1A, LDHA, ENO2, TPI1), and significantly divided patients of HCC into high- and low-risk groups according to overall survival (OS) ($P < 0.001$). We got the Area Under Curve(AUC) value of risk score calculated by PP was 0.777, which much bigger than other clinical parameters. Besides, PP was verified as an independent prognosis-related parameter (in univariate analyse, $HR=1.484$, 95% $CI=1.342-1.642$, $P < 0.001$; in multivariate analyse, $HR=1.414$, 95% $CI=1.258-1.588$, $P < 0.001$). Finally, the application of PP in clinic was concluded that the higher the patient's risk score, the higher the corresponding tumor stage and T stage, and the patient's prognosis was poor.

Conclusions This study provides hypoxic related molecular targets for the therapeutic intervention. In addition, an individualized prognostic predictor was constructed to predict prognosis for HCC patients .

Background

Accumulating evidence has revealed that above 50% of locally advanced solid tumors may show hypoxic tissue regions. The distribution of these areas in tumors is heterogeneous[1]. Hypoxia closely relates to malignant progression and seems to be a sign of poor prognosis in almost all solid tumors, especially in 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]. Sustained effect of hypoxia may also lead to clinically more aggressive phenotypes[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 every year[5, 6]. Thus it is predicted that liver cancer may become the sixth most common 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 related to the prognosis of HCC. The 5-year survival rate for HCC patients was approximately less than 18%[8, 9]. Therefore, finding the key molecular biomarkers focused on hypoxia which 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 measures against HCC.

In recent years, progress using high-throughput sequencing technologies has greatly expanded the research of cancer genome. Herein, we collected 349 HCC patients information from The Cancer Genome Atlas (TCGA)[10]. We compared the expression profile of hypoxic related genes (HRGs) between HCC and non-tumor samples by applying the Limma package in R statistical software. These genes were named, classified and localized through bioinformatic. We also tried to explore their relationship with signaling pathways. Furthermore, in order to make better use of the complementary value between genes and clinic characteristics, 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) outcome of patients. These findings provide new research targets and therapeutic strategies for HCC patients, and provide a reliable theoretical basis for judging treatment outcomes and assessing prognosis.

Results

Differentially expressed HRGs

A total of 374 HCC tissue samples and 50 non-tumor specimens were obtained for RNA-seq and clinical data. from TCGA. By screening patients with clinical follow-up for more than one month, our current study involved 349 patients with primary HCC. A list of 200 HRGs involved in hypoxic regulatory pathways reported in the literature were 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 up-regulated and 9 were down-regulated. (Figure 1A, B). Then, we visualized the expression patterns of these HRGs between HCC and normal tissues by making boxplot (Figure 1C). The boxplot showed expression patterns that 28 up-regulated genes (ALDOA, ANKZF1, ANXA2, B4GALNT2, BCAN, CA12, COL5A1, DDIT3, DPYSL4, DTNA, EFNA3, GPC3, HOXB9, INHA, KDELR3, KIF5A, LOX, NDRG1, P4HA2, PDGFB, PFKF, PGF, PPFIA4, RRAGD, SLC2A1, SLC2A5, STC1, and STC2) and 9 down-regulated genes (DCN, FBP1, FOS, MT1E, MT2A, PCK1, PLAC8, SERPINE1, and ZFP36).

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 and KEGG pathway enrichment of these genes were reviewed. respectively Figure 2 and Figure 3. We found the results that the top three GO terms for biological processes (BP) were: carbohydrate biosynthetic process, glucose metabolic process, and pyruvate metabolic process. The results of top three cellular components (CC) were: endoplasmic reticulum lumen, secretory granule lumen, and cytoplasmic vesicle lumen. The top three molecular function (MF), genes were mostly enriched in terms of protein heterodimerization activity, cell adhesion molecule binding, and receptor ligand activity. The detailed results of all the above gene enrichment showed in Figure 2 A-D. Through the enrichment analysis function in the KEGG pathway,, we found many important pathways associated with these genes such as HIF-1 signaling pathway,

Gluconeogenesis, Carbon metabolism, MAPK signaling pathway , PI3K-Akt signaling pathway, and so on (Figure 3 A-D). As shown in the circle plot of Figure 3C, the change from the Z-score indicated mostly related signaling pathways were more inclined to be increased.

Identification of prognostic HRGs

Prognostic HRGs were computed with univariate Cox regression model by SPSS 22.0. The relationship between the expression 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, a optimal equation for the prognostic predictor (PP) was further conducted by multivariate Cox proportional hazards regression model. Finally, 10 genes (HDLBP, SAP30, PFKP, DPYSL4, SLC2A1, PGK1, ERO1A, LDHA, ENO2, TPI1) were obtained and incorporated into the final model. Among these genes, 8 genes (HDLBP, SAP30, PFKP, DPYSL4, SLC2A1, ERO1A, LDHA, TPI1) were risk factors for HCC patients, 2 genes (PGK1 and ENO2) were protective factors (Table 2).

Construction and definition of the PP

In the multivariate Cox proportional hazards regression model, Y (survival time and status) is obtained by multiplying the contribution weight of every gene and X (expression value of each gene) and then adding them. The formula of PP is expressed as follows: $PP = (0.5003 \times \text{expression value of HDLBP}) + (0.4192 \times \text{expression value of SAP30}) + (0.2345 \times \text{expression value of PFKP}) + (0.3718 \times \text{expression value of DPYSL4}) + (0.4166 \times \text{expression value of SLC2A1}) + (-0.4244 \times \text{expression value of PGK1}) + (0.3085 \times \text{expression value of ERO1A}) + (0.5351 \times \text{expression value of LDHA}) + (-0.5150 \times \text{expression value of ENO2}) + (0.3658 \times \text{expression value of TPI1})$. It is worth noting that the coefficients of the genes were positive, the expression of these genes were beneficial to increase the OS of HCC patients. Conversely, the negative coefficient of the gene means that these genes may shorten the OS of patients with HCC.

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 5A). 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 ($P < 0.001$, Figure 5B). Besides, the high-risk group had more deaths than the low-risk group over time (Figure 5C). All these results indicated that the prediction ability of the equation was very

well. Figures 5D-E showed the distribution of patients according to their prognostic risk. We could see intuitively that the high risk group had a much higher number of deaths than the low risk group.

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 integrated age, gender, tumor grade, AJCC TNM stage with risk score. We used the time-dependent ROC curve analysis as the standard to determine the prediction effect of PP on the prognosis of HCC patients. We got the AUC value of risk score was 0.777, which much bigger than other clinical parameters (Figure 6A). Thus, compared with other indicators, PP had important reference value in predicting patients' prognosis.

Furthermore, risk score remained as an independent prognostic indicator for HCC patients in univariate and multivariate analyses, after adjusting for clinicopathological features such as age, gender, tumor grade, AJCC TNM stage. (HR=1.484, 95% CI=1.342–1.642, $P<0.001$, Figure 6B and HR=1.414, 95% CI=1.258–1.588, $P<0.001$, Figure 6C).

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

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[11]. However, incidence and HCC-specific mortality still continue to increase[12–14]. Although some progress has been made in molecular targeted therapy for HCC, the results are still unsatisfactory[15]. It is necessary 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[16]. In particular, the announcement of the 2019 Nobel Prize in physiology or medicine has increased researchers' enthusiasm for exploring the mechanism of hypoxia in the field of cancer. Increasing evidence indicates that hypoxia closely related to the migration and malignant progression of HCC[8, 17]. However, most studies have focused only on hypoxia by studying signaling genes[18, 19]. 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 that 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 non-tumor tissues. To discover the role of genes in HCC progression, GO and KEGG analysis of the differential expression genes were 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, 20]. 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[21, 22]. 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 survival analysis revealed that 27 HRGs were associated with OS in the TCGA database. Further multivariate survival analysis helped us determine 10 vital prognostic HRGs (HDLBP, SAP30, PFKP, DPYSL4, SLC2A1, PGK1, ERO1A, LDHA, ENO2, TPI1) to construct of the PP. Considering the potential molecular mechanism of the 10 HRGs, rarely reports of the function and mechanism of PFKP, DPYSL4, SLC2A1, ERO1A or TPI1 have been published in HCC. However, among these 10 HRGs, five of them have been studied, including HDLBP, SAP30 and LDHA were upregulated in HCC and promotes the growth of HCC cells. PGK1 and ENO2 were downregulated which could inhibit the growth of cancer cells.

HDLBP (high density lipoprotein binding protein), a positively regulated gene, its closely related to the cancer process[23]. HDLBP can promotes HCC cell proliferation and tumorigenesis. Consistent with many other studies, HDLBP overexpressed in other types of cancers[24]. Intriguingly, in breast cancer, HDLBP may be a tumor suppressor to accelerate the degradation and inhibit the translation of the c-fms proto-oncogene mRNA[25]. 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[26]. It has been demonstrated that SAP up-regulated in serum samples from HCC patients. 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[27]. GK1 (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[28, 29]. These findings demonstrated a novel role of PGK1 in HCC progression. ENO2 (Enolase 2) encodes an enolase isoenzyme which is regarded as a sensitive and specific biomarker for neuroendocrine tumours[30]. Additionally, ENO2 was believed to be an indicator in

the diagnosis and prognosis rather than a potential target for therapy in renal tumor patients[31]. 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.

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[32, 33]. 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 (Fig. 5A). At the same time, we combined age, gender, tumor grade, AJCC TNM stage with risk score and further analyzed through ROC curve, and found that the AUC value of risk score was the largest (AUC = 0.777, Fig. 6A). Secondly, in the process of univariate and multivariate analysis, by integrating the clinicopathological features, we found that the risk score was still an independent indicator for evaluating the prognosis of patients with HCC. (Fig. 6B, Fig. 6C). 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.

Conclusions

In conclusion, through a comprehensive bioinformatics analysis with HRGs expression profiles and corresponding clinical features, we identified 10 prognostic HRGs and constructed the model of the PP. This comprehensive study of multiple databases helps us gain insight into the biological properties of HCC and provides potential molecular targets for subsequent basic research. The PP may help us to make a better judgment of patients' prognosis, so that we can better intervene in patients' progress.

Methods

Data acquisition

MSigDB, the Molecular Signatures Database (<http://software.broadinstitute.org>

[/gsea/msigdb/index.jsp](http://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[34, 35], 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. 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.

Differentially expressed HRGs enrichment analysis

We screened for differences in HRGs expression between HCC and non-tumor samples by applying the Limma package in R statistical software[36]. The significantly differentially expressed HRGs were defined as the adjusted p value is less than 0.05, the gene expression changes at least 2-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 genes enrichment were visualized by R language clustering package.

Construction of the individualized prognostic predictor model based on HRGs

We extracted the expression profile of HRGs downloaded from TCGA in a 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 get the optimal solution of the model by narrowing down the relevant genes step by step. Finally, the PP was composed of the 10 key genes 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 whether hypoxia related gene model could be used as an independent predictor of prognosis in the TCGA cohort of patients with HCC, 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, 2, 3 and 4), lymph node metastasis (positive or negative) and distant metastasis (positive or negative).

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 analyses were used for illustrating the correlation between genes expression profiles and OS. The Multivariate Cox proportional hazards 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.

Abbreviations

HCC:hepatocellular carcinoma; HRGs:hypoxic related genes; PP:prognostic predictor ; OS:overall survival; AUC:Area Under Curve; TCGA:The Cancer Genome Atlas; MsigDB:the Molecular Signatures Database;

GO:Gene Ontology; KEGG:Kyoto Encyclopedia of Genes and Genomes; K-M:Kaplan-meier; ROC:Receiver operating characteristic curve; BP:biological processes; CC:cellular components; MF:molecular function;

Declarations

Acknowledgements

We are grateful to the patient and his family for their contributions to the study. We would like to thank the TCGA and MSigDB databases for the availability of the data.

Funding

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) .

Availability of data and materials

The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request. The datasets analyzed in this study are available in the TCGA databases (<https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>) and MSigDB (<http://software.broadinstitute.org/gsea/msigdb/index.jsp>).

Authors contribution

HG and JZ designed and conceived the study. YZ, JL, YC and LL searched and collected the majority of the data. HG, ZW and WZ performed the bioinformatics analyses. HG and JZ were major contributors in writing the manuscript. All authors have read and approved the final version of this manuscript.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

¹ Department of Anesthesiology, Shanxi provincial people's Hospital, Taiyuan 030000, China. ² Department of Anesthesiology, Tianjin Children's Hospital, Tianjin 300074 China. ³ Department of Anesthesiology, The First Central Clinical College, Tianjin Medical University, Tianjin 300192, China. ⁴ Department of oncology, Shanxi Province Academy of Traditional Chinese Medicine, Shanxi Province Hospital of Traditional Chinese Medicine, Taiyuan, Shanxi 030012, China

References

1. Vaupel P, Mayer A: **Hypoxia in cancer: significance and impact on clinical outcome.** *Cancer Metastasis Rev* 2007, **26**(2):225-239.
2. Zubiete-Franco I, García-Rodríguez JL, Lopitz-Otsoa F, Serrano-Macia M, Simon J, Fernández-Tussy P, Barbier-Torres L, Fernández-Ramos D, Gutiérrez-de-Juan V, López De Davalillo S *et al*: **SUMOylation regulates LKB1 localization and its oncogenic activity in liver cancer.** *EBioMedicine* 2018, **40**:406-421.
3. Vaupel P, Mayer A, Hockel M: **Tumor hypoxia and malignant progression.** *Methods Enzymol* 2004, **381**:335-354.
4. Vaupel P, Mayer A: **Hypoxia and anemia: effects on tumor biology and treatment resistance.** *Transfus Clin Biol* 2005, **12**(1):5-10.
5. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: **Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.** *CA Cancer J Clin* 2018, **68**(6):394-424.
6. de Martel C, Ferlay J, Franceschi S, Vignat J, Bray F, Forman D, Plummer M: **Global burden of cancers attributable to infections in 2008: a review and synthetic analysis.** *Lancet Oncol* 2012, **13**(6):607-615.
7. Pan H, Fu X, Huang W: **Molecular mechanism of liver cancer**, vol. 11; 2011.
8. Song Z, Liu T, Chen J, Ge C, Zhao F, Zhu M, Chen T, Cui Y, Tian H, Yao M *et al*: **HIF-1 α -induced RIT1 promotes liver cancer growth and metastasis and its deficiency increases sensitivity to sorafenib.** *Cancer Lett* 2019, **460**:96-107.
9. Su YH, Kim AK, Jain S: **Liquid biopsies for hepatocellular carcinoma.** *Transl Res* 2018, **201**:84-97.
10. Ally A, Balasundaram M, Carlsen R, Chuah E, Clarke A, Dhalla N, Holt R, Jones S, Lee D MY, MA M *et al*: **Comprehensive and Integrative Genomic Characterization of Hepatocellular Carcinoma.** *Cell* 2017, **169**(7):1327-1341.e1323.
11. Yang JD, Hainaut P, Gores GJ, Amadou A, Plymoth A, Roberts LR: **A global view of hepatocellular carcinoma: trends, risk, prevention and management.** *Nature reviews Gastroenterology & hepatology* 2019, **16**(10):589-604.
12. Zhu J, Yin T, Xu Y, Lu XJ: **Therapeutics for advanced hepatocellular carcinoma: Recent advances, current dilemma, and future directions.** *J Cell Physiol* 2019, **234**(8):12122-12132.
13. Yarchoan M, Agarwal P, Villanueva A, Rao S, Dawson LA, Llovet JM, Finn RS, Groopman JD, El-Serag HB, Monga SP *et al*: **Recent Developments and Therapeutic Strategies against Hepatocellular**

Carcinoma. *Cancer Res* 2019, **79**(17):4326-4330.

14. Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, Brenner H, Dicker DJ, Chimed-Orchir O, Dandona R, Dandona L *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 Oncol* 2017, **3**(4):524-548.
15. Huang W, Liu J, Yan J, Huang Z, Zhang X, Mao Y, Huang X: **LncRNA LINC00470 promotes proliferation through association with NF45/NF90 complex in hepatocellular carcinoma.** *Hum Cell* 2019.
16. Bowyer C, Lewis AL, Lloyd AW, Phillips GJ, Macfarlane WM: **Hypoxia as a target for drug combination therapy of liver cancer.** *Anticancer Drugs* 2017.
17. Guo LY, Zhu P, Jin XP: **Association between the expression of HIF-1 α and VEGF and prognostic implications in primary liver cancer.** *Genet Mol Res* 2016, **15**(2).
18. Carbajo-Pescador S, Ordoñez R, Benet M, Jover R, García-Palomo A, Mauriz JL, González-Gallego J: **Inhibition of VEGF expression through blockade of Hif1 α and STAT3 signalling mediates the anti-angiogenic effect of melatonin in HepG2 liver cancer cells.** *Br J Cancer* 2013, **109**(1):83-91.
19. Myung SJ, Yoon J-H, Yu SJ: **STAT3 & Cytochrome P450 2C9: A novel signaling pathway in liver cancer stem cells.** *Biomed Pharmacother* 2012, **66**(8):612-616.
20. Lei Y, Hu Q, Gu J: **Expressions of Carbohydrate Response Element Binding Protein and Glucose Transporters in Liver Cancer and Clinical Significance.** *Pathol Oncol Res* 2019.
21. Ning X, Wang Y, Yan W, Li G, Sang N: **Chitin synthesis inhibitors promote liver cancer cell metastasis via interfering with hypoxia-inducible factor 1 α .** *Chemosphere* 2018, **206**:231-237.
22. Hu F, Deng X, Yang X, Jin H, Gu D, Lv X, Wang C, Zhang Y, Huo X, Shen Q *et al*: **Hypoxia upregulates Rab11-family interacting protein 4 through HIF-1 α to promote the metastasis of hepatocellular carcinoma.** *Oncogene* 2015, **34**(49):6007-6017.
23. Zhou W, Zhao L, Yuan H, Xu L, Tan W, Song Y, Fang X: **A new small cell lung cancer biomarker identified by Cell-SELEX generated aptamers.** *Exp Cell Res* 2019, **382**(2):111478.
24. Yang WL, Wei L, Huang WQ, Li R, Shen WY, Liu JY, Xu JM, Li B, Qin Y: **Vigilin is overexpressed in hepatocellular carcinoma and is required for HCC cell proliferation and tumor growth.** *Oncol Rep* 2014, **31**(5):2328-2334.
25. Woo HH, Yi X, Lamb T, Menzl I, Baker T, Shapiro DJ, Chambers SK: **Posttranscriptional suppression of proto-oncogene c-fms expression by vigilin in breast cancer.** *Mol Cell Biol* 2011, **31**(1):215-225.
26. Ferrin G, Ranchal I, Llamaza C, Rodriguez-Peralvarez ML, Romero-Ruiz A, Aguilar-Melero P, Lopez-Cillero P, Briceno J, Muntane J, Montero-Alvarez JL *et al*: **Identification of candidate biomarkers for hepatocellular carcinoma in plasma of HCV-infected cirrhotic patients by 2-D DIGE.** *Liver Int* 2014, **34**(3):438-446.
27. Chung TW, Kim EY, Han CW, Park SY, Jeong MS, Yoon D, Choi HJ, Jin L, Park MJ, Kwon YJ *et al*: **Machilin A Inhibits Tumor Growth and Macrophage M2 Polarization Through the Reduction of Lactic**

- Acid. *Cancers (Basel)* 2019, 11(7).**
28. Xie H, Tong G, Zhang Y, Liang S, Tang K, Yang Q: **PGK1 Drives Hepatocellular Carcinoma Metastasis by Enhancing Metabolic Process.** *Int J Mol Sci* 2017, **18**(8).
 29. Hu H, Zhu W, Qin J, Chen M, Gong L, Li L, Liu X, Tao Y, Yin H, Zhou H *et al*: **Acetylation of PGK1 promotes liver cancer cell proliferation and tumorigenesis.** *Hepatology* 2017, **65**(2):515-528.
 30. Isgrò MA, Bottoni P, Scatena R: **Neuron-Specific Enolase as a Biomarker: Biochemical and Clinical Aspects.** *Adv Exp Med Biol* 2015, **867**:125-143.
 31. Luo T, Chen X, Zeng S, Guan B, Hu B, Meng Y, Liu F, Wong T, Lu Y, Yun C *et al*: **Bioinformatic identification of key genes and analysis of prognostic values in clear cell renal cell carcinoma.** *Oncol Lett* 2018, **16**(2):1747-1757.
 32. Zheng Y, Yu K, Huang C, Liu L, Zhao H, Huo M, Zhang J: **Integrated bioinformatics analysis reveals role of the LINC01093/miR-96-5p/ZFAND5/NF-kappaB signaling axis in hepatocellular carcinoma.** *Exp Ther Med* 2019, **18**(5):3853-3860.
 33. Wan Z, Zhang X, Luo Y, Zhao B: **Identification of Hepatocellular Carcinoma-Related Potential Genes and Pathways Through Bioinformatic-Based Analyses.** *Genet Test Mol Biomarkers* 2019, **23**(11):766-777.
 34. Liberzon A, Birger C, Thorvaldsdottir H, Ghandi M, Mesirov JP, Tamayo P: **The Molecular Signatures Database (MSigDB) hallmark gene set collection.** *Cell Syst* 2015, **1**(6):417-425.
 35. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES *et al*: **Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.** *Proc Natl Acad Sci U S A* 2005, **102**(43):15545-15550.
 36. Yu G, Wang LG, Han Y, He QY: **clusterProfiler: an R package for comparing biological themes among gene clusters.** *OMICS* 2012, **16**(5):284-287.

Tables

Table 1 The associated HRGs with OS for univariate Cox regression model in HCCpatients

ID	HR	HR.95%CI	P-value
HDLBP	1.76065475554956	1.1913222761572- 2.60207101829618	0.00453466972203786
SAP30	1.91959358205097	1.46522457295881-2.51486331055061	2.22390390879862e-06
NDRG1	1.34469639494993	1.17281837769284-1.54176335311904	2.19027721354174e-05
ADORA2B	1.55966890253061	1.16388339165171-2.09004364437993	0.00291894799506454
PFKP	1.26452080041416	1.1271911886144-1.41858175510195	6.30340045691965e-05
DPYSL4	1.54885483238764	1.11972059918072-2.14245526389871	0.00821531787802844
PFKFB3	1.20837800501272	1.06584452549981-1.36997223146949	0.00311915122509499
SLC2A1	1.62046847152965	1.38336161732254-1.89821521310099	2.2248239035816e-09
VEGFA	1.43129203349915	1.14062682980327-1.79602726468522	0.00196126023621722
HMOX1	1.25919849013278	1.09400331393543-1.4493382399811	0.00131766847435219
SDC3	1.37383942166371	1.10908277858783-1.7017979117127	0.00363911893920355
ATP7A	2.01311206689882	1.28397811820806-3.15630004625743	0.0022933443503931
PGK1	1.44610164059818	1.16681166410303-1.79224292940911	0.000754424227811794
ERO1A	1.48146745049227	1.17333443232747-1.87052024248063	0.00095472636841490
XPNPEP1	2.02964243992359	1.31993211590091-3.12095477056204	0.00126250541602928
LDHA	2.08667811334332	1.59510255119679-2.72974646391176	8.01704374515246e-08
ENO2	1.29276224345037	1.09688339722492-1.52362067136672	0.0021910543252294
GRHPR	0.73434761865719	0.595415431105967-0.905697764711638	0.00390630491639015
UGP2	0.74052056373479	0.595909710542951-0.92022448302522	0.00672918301280524
SLC6A6	1.28627183707835	1.08185403431661-1.52931466388258	0.00435902717618689
MAP3K1	1.45729575597612	1.10637104849598 -1.91952864572245	0.00738173976636133
BNIP3L	1.429262964912	1.12032257711033-1.82339681856456	0.00404897204228161
TPI1	1.6050022442635	1.23862141303868-2.07975752475583	0.00034544276992469
RRAGD	1.31789370530495	1.07275132642648-1.61905539121368	0.0085699110893533
TES	1.28464795991929	1.07297845423873-1.53807410988107	0.00639618215197407
SERPINE1	1.15781169644034	1.04780421344428-1.27936870954886	0.00401837021242022
JMJD6	1.76950202770856	1.31928911762902-2.37335196980315	0.00013913471553068

Abbreviations: HR, hazard ratio; C,I, confidence interval.

Table 2 The associated HRGs with OS for multivariate Cox proportional hazards regression model

ID	Coefficient	HR	HR.95%CI	P-value
HDLBP	0.500288	1.649195	1.0947-2.484557	0.016725
SAP30	0.419237	1.5208	1.098428-2.105586	0.011553
PFKP	0.234499	1.264275	1.044329-1.530545	0.016183
DPYSL4	0.371773	1.450304	1.02894-2.04422	0.033765
SLC2A1	0.416578	1.516763	1.119095-2.055741	0.007247
PGK1	-0.42436	0.654191	0.463678-0.92298	0.015678
ERO1A	0.308536	1.36143	0.96342-1.923867	0.080335
LDHA	0.535135	1.707679	1.229596-2.371647	0.001406
ENO2	-0.515	0.5975	0.426939-0.8362	0.002672
TPI1	0.36584	1.441725	1.019581-2.03865	0.038483

Abbreviations: HR, hazard ratio; C,I, confidence interval.

Figures



Figure 1

Differentially expressed hypoxic related genes (HRGs) between hepatocellular carcinoma (HCC) and normal liver 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. Blue shows no difference between HCC and normal liver tissues. (B) Hierarchical clustering of differentially expressed HRGs expression levels. (C) The expression patterns of Box plots displayed 37 differentially expressed HRGs in HCC types and paired non-tumor samples. Each red dot represents a distinct tumor sample and green a normal sample. The red bar above the gene name shows a significantly high expression and the green bar a low expression.



Figure 2

The gene ontology (GO) terms function enrichment of differentially expressed hypoxic related genes (HRGs). (A) The bar plot of enriched GO terms describe these genes from three levels. (B) The bubble plot of GO is assigned as in the barplot (the higher the more significant). (C) The circle plot of GO shows the top 10 enriched GO terms for BP, CC, and MF. The outer circle shows a scatter plot for each term of the logFC of the assigned genes. Red circles display up-regulation, and blue ones down-regulation. (D) The heatmap of the relationship between HRGs and GO terms.



Figure 3

Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis of differentially expressed hypoxic related genes (HRGs). (A) The bar plot shows the top enriched KEGG pathways. (B) The bubble plot of KEGG is assigned as in the barplot (the higher the more significant). (C) The circle plot of KEGG shows the top 10 enriched KEGG pathways. The outer circle shows a scatter plot for each pathway of the logFC of the assigned genes. Red circles display up-regulation, and blue ones down-regulation. (D) The heatmap of the relationship between HRGs and pathways. The color of each block is associated with the logFC values.



Figure 4

The relationships between the expression profiles of related genes (HRGs) and OS. In these 37 HRGs, Red squares display risk factors, and the other 2 green squares display protective factors.



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 number of patients in different riskgroups. (D) The overall survival of patients in the high- and low-risk group. (E) The heatmap shows the distribution of these 10 key genes expression profiles in the TCGA dataset.



Figure 6

The accuracy of hypoxia-related prognostic predictor (PP) in hepatocellular carcinoma patients. (A) The risk score obtained the largest AUC through the time-dependent ROC curve analysis. (B) Risk score with clinicopathological features in univariate analyse. (C) Risk score with clinicopathological features in multivariate analyse. 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.