Proceedings of ICBioMed 2025 Symposium: Computational Modelling and Simulation for Biology and Medicine
DOI: 10.54254/2753-8818/2025.LD27328

Identification and Prognostic Evaluation of Biomarker Genes in Hepatocellular Carcinoma: Focus on EZH2 as a Potential Biomarker and Therapeutic Target

Rui Jin

Ulink College of Shanghai, Shanghai, China
jinrui0116@163.com

Abstract. Hepatocellular carcinoma (HCC) ranks among the most prevalent and aggressive cancers globally, characterized by high rates of illness and death. Early diagnosis and effective treatment are hindered by a lack of reliable biomarkers and a comprehensive understanding of its molecular mechanisms. We carried out a bioinformatics analysis in this study, employing the TCGA and GEO databases to identify genes with varying expression levels between HCC and normal liver tissues. The TCGA dataset identified 395 genes that were upregulated and 631 that were downregulated, whereas the GSE45267 dataset found 2,452 upregulated genes and 446 downregulated ones. These DEGs were primarily involved in critical biological processes such as cell cycle regulation, DNA replication, and metabolism. Gene set enrichment analysis (GSEA) indicated that pathways associated with mitotic division and cell cycle progression were enriched with upregulated genes, whereas downregulated genes were linked to immune response and metabolic pathways. Further analysis of gene interaction networks revealed several key hub genes, including EZH2, MCM2, and ALDH2, which were associated with poor and favorable prognosis, respectively. Survival analysis identified EZH2, GINS1, and MCM2 as adverse prognostic markers and ALDH2, ADH4, and PON1 as favorable prognostic markers. A Lasso Cox regression model was developed, incorporating these genes to construct a prognostic risk score, which showed high predictive accuracy for patient survival. These findings offer insights into the molecular landscape of HCC, identify potential diagnostic and prognostic biomarkers, and propose novel therapeutic targets for personalized treatment strategies. Future studies should validate these biomarkers and assess their clinical utility in HCC management.

Keywords: TCGA Database DNA Replication, Prognostic Model Hepatocellular Carcinoma, Differentially Expressed Genes

1. Introduction

Hepatocellular carcinoma (HCC) is the leading type of primary liver cancer and is a significant contributor to cancer-related mortality globally. It is closely linked to chronic liver conditions, including infections from hepatitis B virus (HBV) and hepatitis C virus (HCV), cirrhosis, and non-

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2025.LD27328

alcoholic fatty liver disease (NAFLD) [1,2]. Even with improvements in diagnostic methods and treatments, the outlook for HCC is still bleak because it is usually diagnosed at a late stage and is resistant to standard therapies. Early detection and reliable prognostic markers are essential for improving clinical outcomes, as they would allow for better patient stratification, timely intervention, and more personalized treatment strategies [3-5].

The molecular mechanisms driving HCC are intricate and multifaceted, involving dysregulation of key signaling pathways that control cell cycle progression, apoptosis, and cellular metabolism. These include alterations in the PI3K/Akt pathway, the MAPK pathway, and cell cycle regulators, among others. However, the precise molecular changes that underlie HCC pathogenesis remain incompletely understood, and new biomarkers and therapeutic targets need to be discovered [6-8].

One effective strategy for discovering molecular changes in HCC is through the study of differentially expressed genes (DEGs). The advent of high-throughput sequencing technologies has enabled the large-scale identification of DEGs in various cancers, including HCC. These technologies allow for comprehensive profiling of gene expression patterns, providing insights into the dysregulated pathways and molecular signatures that drive tumorigenesis [9]. In HCC, the most frequently identified DEGs are those that regulate the cell cycle, DNA replication, and mitotic division. For instance, genes such as BIRC5, CDC20, and CCNB1 have been found to be upregulated in HCC and are associated with tumor progression, while other genes involved in immune responses and metabolic processes are often downregulated in HCC tissues [10].

To investigate these gene expression changes in HCC, we have leveraged data from well-established datasets such as The Cancer Genome Atlas (TCGA) [11] and Gene Expression Omnibus (GEO) [12]. The TCGA dataset provides large-scale genomic data from numerous cancer types, including HCC, while the GEO dataset is a valuable resource for gene expression studies. By performing a comprehensive bioinformatics analysis of these datasets, we aim to identify DEGs between HCC and normal liver tissues and analyze their associated biological pathways.

In our study, we utilized both the TCGA and GSE45267 datasets to identify and categorize upregulated and downregulated genes in HCC. The TCGA dataset revealed 395 upregulated genes and 631 downregulated genes, with key genes such as BIRC5, CDC20, and PTTG1 being upregulated, and others like CLEC4M, GDF2, and FCN2 being downregulated. Further analysis of the GSE45267 dataset confirmed the presence of 2,452 upregulated genes and 446 downregulated genes, with significant overlap in pathways such as mitotic division, DNA replication, and cellular transport. Using the STRING database, we examined gene interaction networks and carried out survival analysis to determine the prognostic value of these DEGs. In particular, we identified genes such as EZH2, GINS1, and MCM2 as potential adverse prognostic markers, while genes like ALDH2, ADH4, and PON1 were associated with favorable outcomes.

2. Methods

2.1. Differential expression analysis

We obtained STAR-counts data and related clinical details for HCC from the TCGA database (https://portal.gdc.cancer.gov). The data were then processed to extract TPM (Transcripts Per Million) format data, followed by log2(TPM+1) normalization. After selecting samples with both RNAseq data and clinical details, 632 samples were picked for additional analysis.

To identify differentially expressed genes (DEGs), we employed bioinformatics methods implemented in R. Using the limma package in R, the DEGs between HCC and normal liver tissues

DOI: 10.54254/2753-8818/2025.LD27328

were identified. Genes with adjusted p-values less than 0.05 and a log2 fold change greater than 1 were deemed significantly differentially expressed.

2.2. Gene interaction networks

The STRING database was employed to investigate the interaction networks of DEGs by examining protein-protein interactions. We used the STRING web tool to build interaction networks for genes that were either upregulated or downregulated. The protein interactions were visualized using Cytoscape (version 3.8.2). We then applied DBSCAN clustering to identify potential functional clusters within the gene interaction networks.

2.3. Prognostic analysis

To assess the prognostic significance of DEGs in HCC, survival analysis was conducted. Kaplan-Meier survival plots were created for each gene, and the log-rank test compared survival rates between groups with high and low expression. Univariate Cox proportional hazards regression models were used to calculate hazard ratios (HRs) with 95% confidence intervals (CIs). Significant prognostic factors were further assessed using multivariate Cox regression analysis to account for potential confounding clinical factors.

For constructing a nomogram model, we selected key genes identified in the survival analysis and combined them with relevant clinical parameters. Using the rms package in R, a nomogram was created, and calibration curves were drawn to evaluate the model's precision in forecasting survival probabilities at 1, 3, and 5 years.

2.4. Lasso Cox regression model

We employed Lasso Cox regression to pinpoint crucial genes linked to survival outcomes, aiming to build a strong prognostic model. The Lasso technique was executed using the glmnet package in R. We selected features with the optimal lambda parameter.

Patients were divided into high- and low-risk categories based on the risk score. Kaplan-Meier survival curves and log-rank tests were utilized to assess survival differences between these groups.

2.5. Statistical analysis

R software was used for all statistical analyses, and a p-value below 0.05 was deemed statistically significant. Univariate and multivariate Cox regression models were used to evaluate prognostic factors, with β-values (Cox coefficients) indicating the link between each variable and patient survival. The significance of each variable was assessed using the Wald test, where higher Wald values suggest stronger associations. Standard errors (SE) were calculated for each β-value to quantify the uncertainty of the model's predictions.

3. Results

3.1. DEGs between HCC and normal liver tissue

This research focused on finding DEGs between HCC and normal liver tissues. The TCGA dataset and the GSE45267 dataset from the NCBI were selected for analysis, both relevant to HCC. Bioinformatics and R-based analysis were employed to filter out the DEGs. From the TCGA dataset,

DOI: 10.54254/2753-8818/2025.LD27328

395 upregulated genes and 631 downregulated genes were identified (Figure 1A). Notably, genes such as BIRC5, CDC20, PTTG1, CCNB1, and UBE2C were upregulated in HCC, while CLEC4M, GDF2, FCN2, CRHBP, COLEC10, and FCN3 were downregulated (Figure 1B). Upregulated genes were primarily involved in pathways such as mitotic nuclear division, chromosome segregation, the cell cycle, nucleocytoplasmic transport, DNA replication, ribosome biogenesis, and spliceosome function. Conversely, the downregulated genes were found to be enriched in pathways associated with retinol metabolism, and chemical carcinogenesis, complement and coagulation cascades, drug metabolism, organic acid catabolism, xenobiotic metabolism, and fatty acid catabolism (Figure 1C).

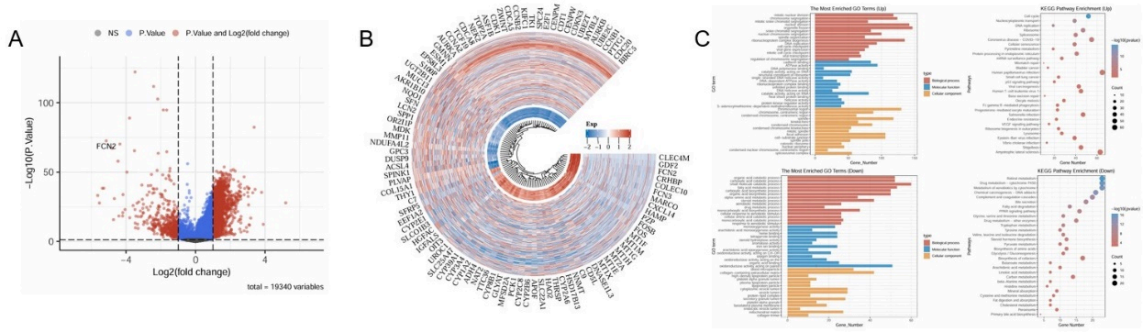


Figure 1. Differentially expressed genes (DEGs) in the TCGA dataset. This figure presents the identification of DEGs between hepatocellular carcinoma (HCC) and normal liver tissues. A total of 395 genes were upregulated and 631 genes were downregulated. The top 10 most significantly upregulated and downregulated genes, such as BIRC5, CDC20, PTTG1, UBE2C (upregulated) and CLEC4M, GDF2, FCN2, CRHBP (downregulated), are displayed. Gene ontology and pathway enrichment analysis revealed that the upregulated genes are primarily involved in mitotic nuclear division, the cell cycle, DNA replication, while downregulated genes are enriched in pathways related to retinol metabolism, drug metabolism, and fatty acid catabolism

Further analysis of the GSE45267 dataset from the GEO database revealed 2,452 upregulated genes and 446 downregulated genes (Figure 2A). Notably, genes such as TTK, CENPW, NCAPG, BUB1B, KIF4A, and CCNB1 were upregulated in HCC, while IDO2, AOPF, NAT2, LPA, C3ORF85, BCO2, and C6 were downregulated (Figure 2B). Upregulated genes in the GEO dataset were predominantly involved in pathways such as mitotic nuclear division, chromosome segregation, the cell cycle, p53 signaling, DNA replication, and mismatch repair. Conversely, downregulated genes were enriched in pathways associated with complement and coagulation cascades, drug metabolism, chemical carcinogenesis, metabolism of xenobiotics, organic acid catabolic process, retinol metabolism, and fatty acid catabolism (Figure 2C).

DOI: 10.54254/2753-8818/2025.LD27328

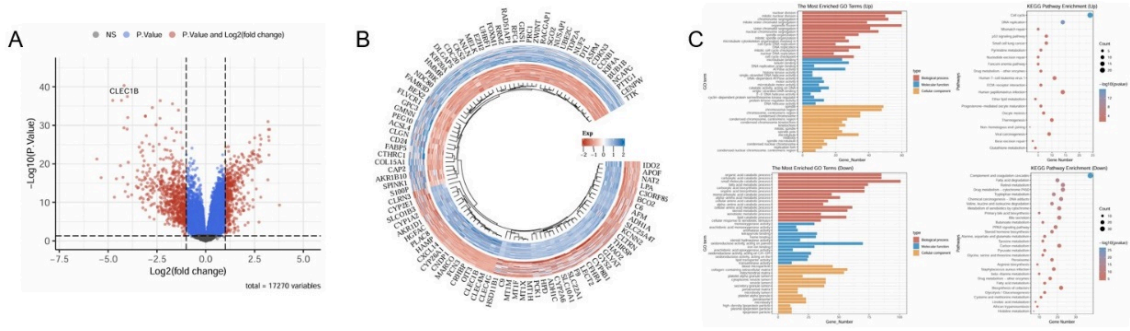


Figure 2. Differentially expressed genes (DEGs) in the GSE45267 dataset. This figure shows the identification of DEGs from the GSE45267 dataset between HCC and normal liver tissues. A total of 2,452 genes were upregulated and 446 genes were downregulated. The top 10 most significantly upregulated genes, such as TTK, CENPW, NCAPG, BUB1B (upregulated), and IDO2, AOPF, NAT2, C6 (downregulated) are presented. Gene ontology and pathway enrichment analysis highlighted that upregulated genes are primarily involved in mitotic nuclear division, the cell cycle, DNA replication, while downregulated genes are associated with retinol metabolism, drug metabolism, and fatty acid catabolism

3.2. Gene interaction networks in hepatocellular carcinoma and normal liver tissue

The intersection of DEGs from the TCGA and GEO databases revealed 324 co- upregulated genes (Figure 3A) and 321 co-downregulated genes (Figure 3B). These common upregulated and downregulated genes were further analyzed for their interaction networks using the STRING database, revealing widespread gene interactions. The DBSCAN clustering algorithm identified 12 clusters for upregulated genes and 24 clusters for downregulated genes. The most central cluster in the upregulated genes was designated as cluster 1, serving as a hub gene, while the most central clusters in the downregulated genes were clusters 1 and 2, which were also considered hub genes. Figures 4 and 5 illustrate the clustering of interaction proteins and their network distributions. A total of 94 hub genes showed increased expression, while 37 showed decreased expression.

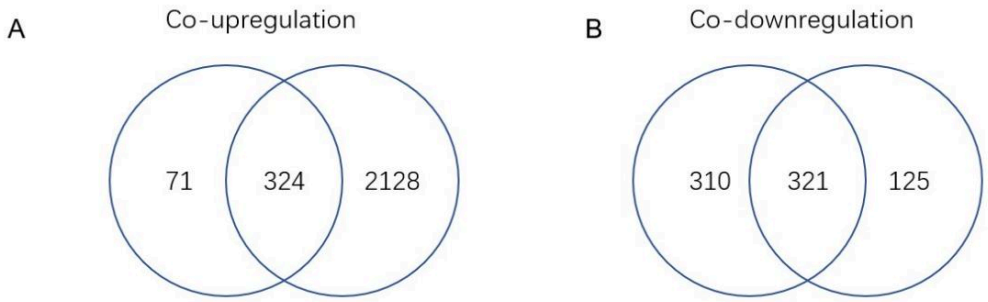


Figure 3. Intersection of DEGs from the TCGA and GEO datasets. This Venn diagram shows the overlap between the DEGs identified from the TCGA and GSE45267 datasets. A total of 324 common upregulated genes and 321 common downregulated genes were identified, underscoring the consistency between the two datasets in terms of differential gene expression in HCC

DOI: 10.54254/2753-8818/2025.LD27328

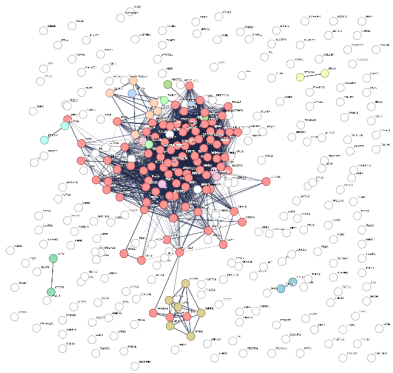


Figure 4. Gene interaction network of upregulated genes. This network graph was constructed using the STRING database and shows the protein-protein interactions (PPIs) for the 324 common upregulated genes identified in both TCGA and GEO datasets. DBSCAN clustering identified 12 clusters for upregulated genes, with cluster 1 as the most central cluster. This network highlights key interactions among upregulated genes, providing insight into their potential functional relationships in HCC

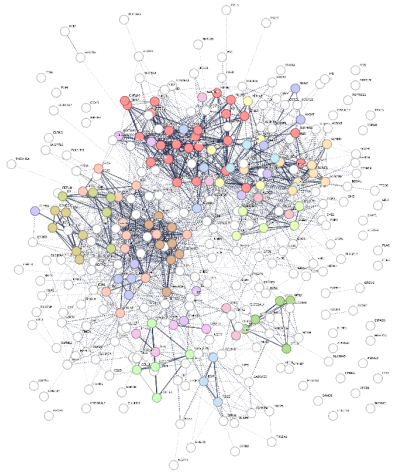


Figure 5. Gene interaction network of downregulated genes. A network graph showing the protein-protein interactions (PPIs) for the 321 common downregulated genes in both TCGA and GEO datasets. DBSCAN clustering identified 24 clusters, with clusters 1 and 2 as the most central clusters. These interactions shed light on the role of downregulated genes in immune responses and metabolic processes in HCC

3.3. Prognostic value of screening genes for hepatocellular carcinoma

We performed survival analysis on the DEGs, indicated that the majority of these genes were associated with poor prognosis. Specifically, EZH2, GINS1, and MCM2 (Figure 6) were found to be adverse prognostic markers even after adjusting for covariates, while ALDH2, ADH4, CYP2C9, APOF, and PON1 were identified as favorable prognostic factors after covariate adjustment (Figure 7). Using both univariate and multivariate analyses, we identified variables suitable for inclusion in a Nomogram. If a gene displayed significant differences in both univariate and multivariate analyses, it was regarded as independent of other clinical factors. The Nomogram model

DOI: 10.54254/2753-8818/2025.LD27328

demonstrated good predictive performance, as evidenced by its close alignment with the calibration curve.

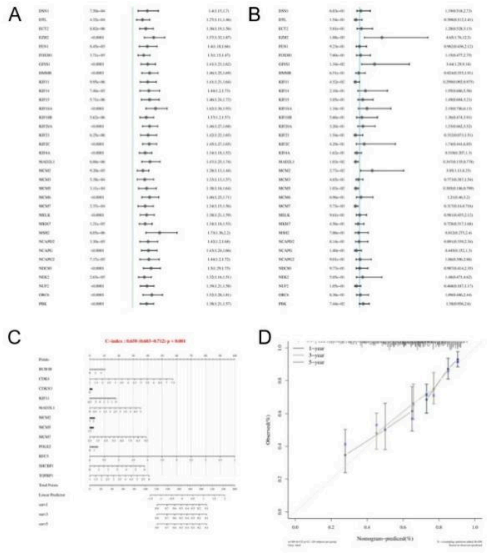


Figure 6. Survival analysis of key DEGs in HCC. Kaplan-Meier survival curves show the prognostic significance of key DEGs, including EZH2, GINS1, and MCM2. These genes were identified as adverse prognostic markers for HCC, with log-rank tests revealing statistically significant survival differences between high and low expression groups. This figure emphasizes the relevance of these genes in predicting patient outcomes

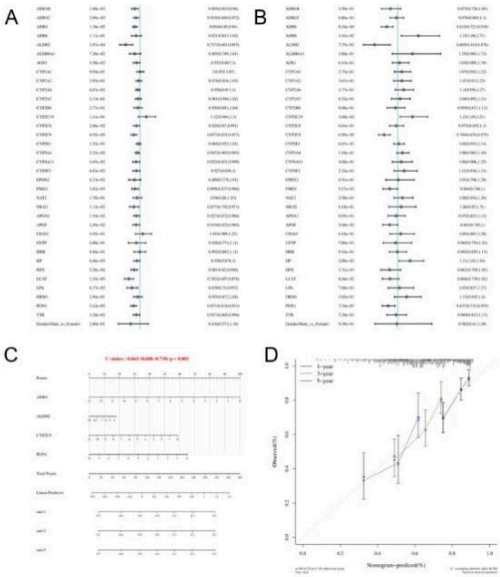


Figure 7. Survival analysis of favorable prognostic genes. Kaplan-Meier survival curves demonstrate the association between favorable prognostic genes (ALDH2, ADH4, PON1) and improved survival in HCC. Statistically significant differences between high and low expression groups were observed, highlighting the potential of these genes as protective biomarkers in HCC prognosis.

DOI: 10.54254/2753-8818/2025.LD27328

In the Nomogram plot, a line segment with a scale represents each variable, indicating the range of possible values for that variable, while the segment's length reflects how much that factor influences the prognostic outcome. On the 'Point' axis, you can see the score for each variable, and the 'Total Point' axis indicates the overall score obtained by adding up the individual scores, with satisfactory predictive accuracy.

3.4. Construction of a prognostic model using Lasso Cox regression

We used Lasso Cox regression to construct a prognostic model based on selected features. The coefficients of these features were determined by the lambda parameter, with the x-axis representing the lambda values and the y-axis representing the coefficients of the independent variables. The relationship between the partial likelihood deviance and log(λ) was depicted in the Lasso Cox regression plot (Figure 8 A&B). The formula for the final risk score is:

Riskscore=(0.2155)×EZH2+(0.0222)×GINS1+(0.0538)×MCM2+(−0.0376)×ADH4+(−0.0503)×CYP2C9+(0.0049)×APOF+(−0.0593)×PON1

The performance of this model was evaluated using Kaplan-Meier (KM) survival curves, where the risk groups were compared using the log-rank test. A hazard ratio (HR) greater than 1 indicated that the model was a risk factor, while an HR less than 1 indicated a protective model. The median survival time and 95% CI were also determined for groups classified as high-risk and low-risk. According to the ROC curve analysis, the model exhibited significant predictive strength, with the AUC value reflecting robust predictive performance. Our risk score model presents the model's effectiveness (Figure 8C).

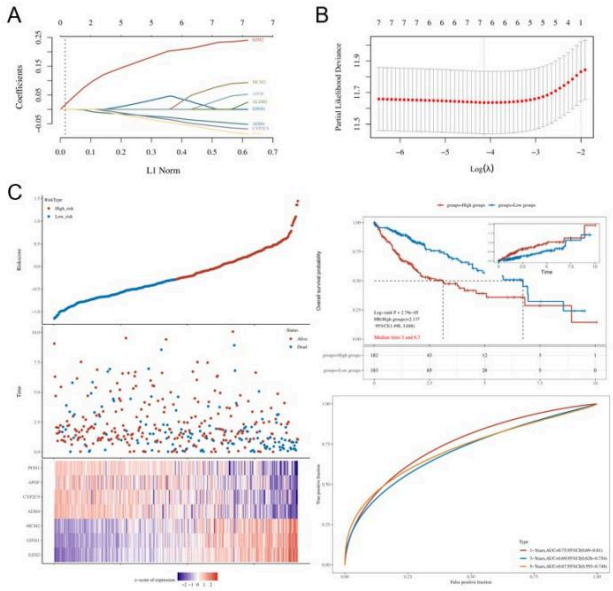


Figure 8. Lasso Cox regression analysis and risk score model construction. Panel A illustrates the Lasso Cox regression analysis used to select key prognostic genes, with the relationship between the lambda parameter and the coefficients of the genes. Panel B displays the formula for the risk score calculation based on the Lasso Cox regression model, incorporating genes such as EZH2, GINS1, MCM2, ADH4, CYP2C9, APOF, and PON1. Panel C shows the receiver operating characteristic (ROC) curve analysis, demonstrating the model's predictive accuracy with a high area under the curve (AUC) value.

DOI: 10.54254/2753-8818/2025.LD27328

3.5. EZH2 expression and its role in tumor stemness

HCC patients were classified into high and low EZH2 expression groups to study the role of the key gene EZH2 in tumor stemness. There was a notable positive correlation between EZH2 expression and tumor stemness, according to Spearman correlation analysis (Figure 9). The correlation coefficient and p-value are displayed, indicating a strong association between EZH2 expression and stemness scores.

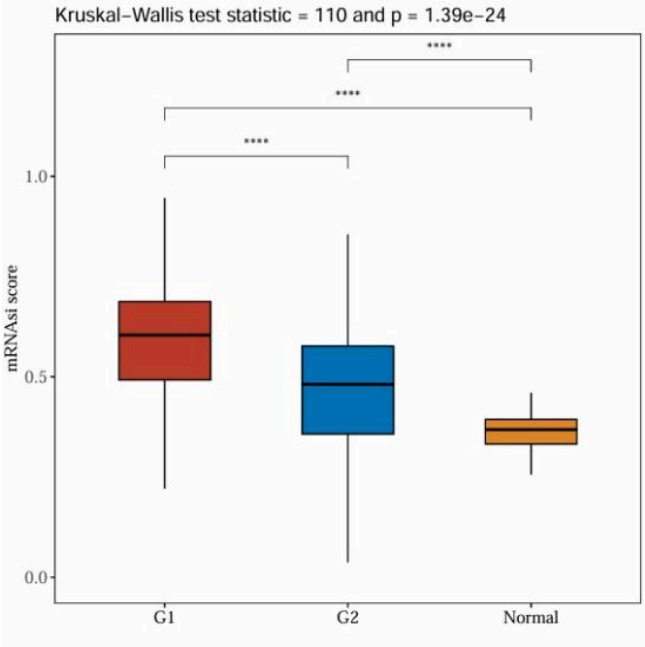
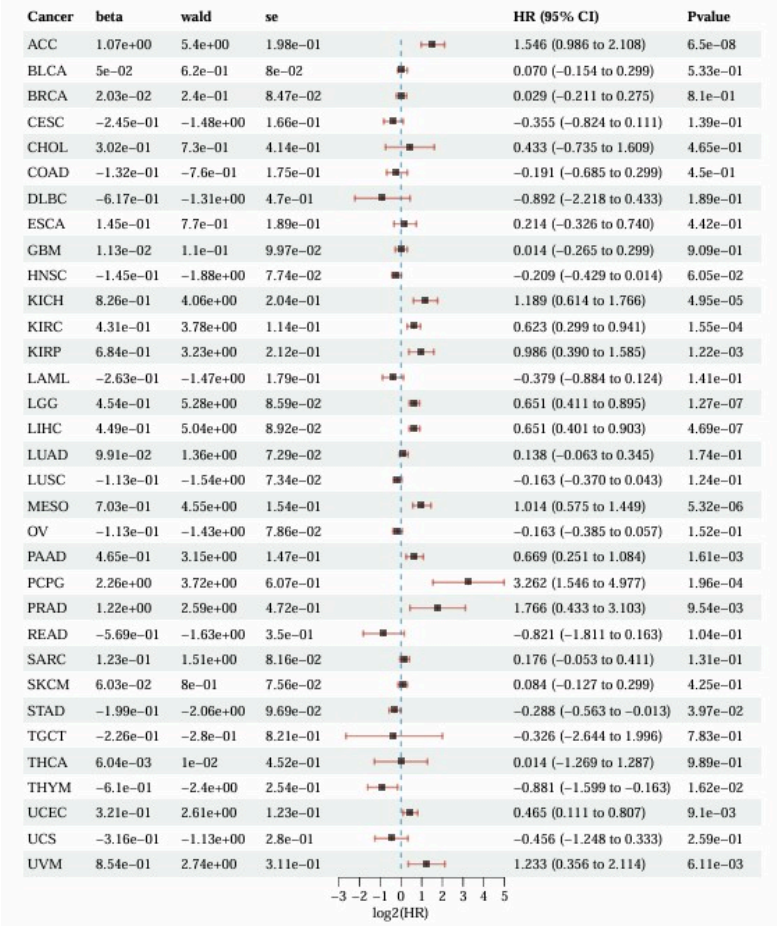


Figure 9. Correlation between EZH2 expression and tumor stemness. This scatter plot illustrates the significant positive correlation between EZH2 expression and tumor stemness scores in HCC patients, with the Spearman correlation coefficient and p-value displayed. This figure suggests that EZH2 may play a critical role in regulating tumor stemness in HCC.

3.6. EZH2 expression across multiple cancer types

To determine the prognostic importance of EZH2 in multiple cancers, we analyzed TCGA data to measure EZH2 mRNA expression in diverse tumor tissues. A total of 36 cancer types and 10,228 samples were analyzed. EZH2 mRNA expression was found to influence prognosis in several cancers, including ACC, KIRP, KICH, LGG, PADD, KIRC, PRAD, UCEC, PCPG, LIHC and UVM (Figure 10). A forest plot, which presents the results of univariate Cox regression for EZH2 in multiple cancers, showed significant associations with p-values, hazard ratios (HR), confidence intervals (CIs), and β-values. These findings suggest that high EZH2 expression is a common feature across various tumors, highlighting its potential as a valuable therapeutic target.

DOI: 10.54254/2753-8818/2025.LD27328



Cancer	beta	wald	se	HR (95% CI)	Pvalue
ACC	1.07e+00	5.4e+00	1.98e−01	1.546 (0.986 to 2.108)	6.5e−08
BLCA	5e−02	6.2e−01	8e−02	0.070 (−0.154 to 0.299)	5.33e−01
BRCA	2.03e−02	2.4e−01	8.47e−02	0.029 (−0.211 to 0.275)	8.1e−01
CESC	−2.45e−01	−1.48e+00	1.66e−01	−0.355 (−0.824 to 0.111)	1.39e−01
CHOL	3.02e−01	7.3e−01	4.14e−01	0.433 (−0.735 to 1.609)	4.65e−01
COAD	−1.32e−01	−7.6e−01	1.75e−01	−0.191 (−0.685 to 0.299)	4.5e−01
DLBC	−6.17e−01	−1.31e+00	4.7e−01	−0.892 (−2.218 to 0.433)	1.89e−01
ESCA	1.45e−01	7.7e−01	1.89e−01	0.214 (−0.326 to 0.740)	4.42e−01
GBM	1.13e−02	1.1e−01	9.97e−02	0.014 (−0.265 to 0.299)	9.09e−01
HNSC	−1.45e−01	−1.88e+00	7.74e−02	−0.209 (−0.429 to 0.014)	6.05e−02
KICH	8.26e−01	4.06e+00	2.04e−01	1.189 (0.614 to 1.766)	4.95e−05
KIRC	4.31e−01	3.78e+00	1.14e−01	0.623 (0.299 to 0.941)	1.55e−04
KIRP	6.84e−01	3.23e+00	2.12e−01	0.986 (0.390 to 1.585)	1.22e−03
LAML	−2.63e−01	−1.47e+00	1.79e−01	−0.379 (−0.884 to 0.124)	1.41e−01
LGG	4.54e−01	5.28e+00	8.59e−02	0.651 (0.411 to 0.895)	1.27e−07
LIHC	4.49e−01	5.04e+00	8.92e−02	0.651 (0.401 to 0.903)	4.69e−07
LUAD	9.91e−02	1.36e+00	7.29e−02	0.138 (−0.063 to 0.345)	1.74e−01
LUSC	−1.13e−01	−1.54e+00	7.34e−02	−0.163 (−0.370 to 0.043)	1.24e−01
MESO	7.03e−01	4.55e+00	1.54e−01	1.014 (0.575 to 1.449)	5.32e−06
OV	−1.13e−01	−1.43e+00	7.86e−02	−0.163 (−0.385 to 0.057)	1.52e−01
PAAD	4.65e−01	3.15e+00	1.47e−01	0.669 (0.251 to 1.084)	1.61e−03
PCPG	2.26e+00	3.72e+00	6.07e−01	3.262 (1.546 to 4.977)	1.96e−04
PRAD	1.22e+00	2.59e+00	4.72e−01	1.766 (0.433 to 3.103)	9.54e−03
READ	−5.69e−01	−1.63e+00	3.5e−01	−0.821 (−1.811 to 0.163)	1.04e−01
SARC	1.23e−01	1.51e+00	8.16e−02	0.176 (−0.053 to 0.411)	1.31e−01
SKCM	6.03e−02	8e−01	7.56e−02	0.084 (−0.127 to 0.299)	4.25e−01
STAD	−1.99e−01	−2.06e+00	9.69e−02	−0.288 (−0.563 to −0.013)	3.97e−02
TGCT	−2.26e−01	−2.8e−01	8.21e−01	−0.326 (−2.644 to 1.996)	7.83e−01
THCA	6.04e−03	1e−02	4.52e−01	0.014 (−1.269 to 1.287)	9.89e−01
THYM	−6.1e−01	−2.4e+00	2.54e−01	−0.881 (−1.599 to −0.163)	1.62e−02
UCEC	3.21e−01	2.61e+00	1.23e−01	0.465 (0.111 to 0.807)	9.1e−03
UCS	−3.16e−01	−1.13e+00	2.8e−01	−0.456 (−1.248 to 0.333)	2.59e−01
UVM	8.54e−01	2.74e+00	3.11e−01	1.233 (0.356 to 2.114)	6.11e−03

Figure 10. EZH2 expression across multiple cancer types. A forest plot showing the mRNA expression levels of EZH2 across 36 different cancer types using TCGA data. The plot includes hazard ratios (HRs), p-values, and confidence intervals (CIs) for various cancers. Significant associations between high EZH2 expression and poor prognosis were observed in several cancer types, including hepatocellular carcinoma (LIHC), highlighting its potential as a prognostic biomarker across different malignancies

4. Discussion

HCC is a major global health challenge, with high morbidity and mortality rates, and its molecular mechanisms remain poorly understood. This research involved an extensive examination of DEGs between HCC and normal liver tissues, utilizing two major datasets: TCGA and GSE45267. Our results, which identified critical genes involved in cell cycle regulation, DNA replication, and metabolic pathways, offer valuable insights into HCC pathogenesis. However, it is essential to contextualize these findings by comparing them with other studies in the field and discussing the clinical relevance of our findings in light of current research.

Our analysis identified several well-established upregulated genes in HCC, such as BIRC5, CDC20, and CCNB1. Known for their involvement in cell cycle progression and mitosis, these genes support the notion that abnormal cell division is a key feature of cancer. Indeed, numerous studies highlight the involvement of these genes in the advancement of tumors. BIRC5 (Survivin) is known to prevent apoptosis and enhance cell survival, establishing it as a significant target for cancer treatment, especially in HCC [13-16]. Similarly, CDC20, a key regulator of the anaphase-

DOI: 10.54254/2753-8818/2025.LD27328

promoting complex, is often overexpressed in various cancers, including HCC, and contributes to uncontrolled cell proliferation by facilitating mitotic progression [17-19]. Our findings align with these reports, further validating the importance of these genes in HCC.

On the other hand, downregulated genes such as CLEC4M, GDF2, and FCN2 identified in our study are associated with immune responses and metabolic processes [20-22]. These findings are consistent with the hypothesis that immune evasion and altered metabolism play critical roles in HCC development. CLEC4M, for instance, is involved in pathogen recognition and immune modulation, and its downregulation has been implicated in immune escape mechanisms in cancer [23]. Likewise, GDF2, which is involved in vascular remodeling, has been suggested to act as a tumor suppressor in various cancers, including HCC [24-26]. Our results support these previous findings, indicating that immune dysregulation and metabolic reprogramming are crucial components of HCC pathogenesis.

To explore the molecular interactions underlying HCC, we constructed gene interaction networks using the STRING database and identified several key hub genes, including EZH2, MCM2, and ALDH2. These hub genes are central to cellular processes such as epigenetic regulation, DNA replication, and detoxification, and have been widely studied [27-30].

EZH2, a key component of the polycomb repressive complex 2 (PRC2), was identified as a significant upregulated gene in our analysis, which is consistent with its role in promoting tumorigenesis by silencing tumor suppressor genes [31-33]. Research indicates that EZH2 is overexpressed in various cancers, such as HCC, and inhibiting it could be a potential therapeutic approach. Our findings provide further evidence of EZH2's importance in HCC and suggest that targeting this gene could be beneficial for therapeutic interventions.

The prognostic value of the DEGs identified in this study was further assessed using survival analysis. We found that genes such as EZH2, GINS1, and MCM2 were associated with poor prognosis, while genes like ALDH2, ADH4, and PON1 were identified as favorable prognostic markers. These findings are consistent with previous studies, which have highlighted the role of EZH2 and MCM2 in poor prognosis in HCC. EZH2's association with adverse outcomes has been well-documented, as its overexpression leads to epigenetic silencing of tumor suppressor genes, promoting tumor growth and metastasis. Similarly, MCM2's role in DNA replication and cell cycle progression links its overexpression to aggressive tumor behavior and poor patient outcomes [31-34].

In contrast, genes like ALDH2, which is involved in alcohol metabolism and detoxification, have been shown to have a protective role in liver cancer. ALDH2 is known to detoxify aldehydes generated during alcohol metabolism, and its downregulation has been linked to increased oxidative stress and liver carcinogenesis. Our study supports these findings, suggesting that ALDH2 may serve as a favorable prognostic marker in HCC.5,Conclusion.

Although our findings offer important insights into the molecular mechanisms of HCC, this study has several limitations. The analysis was primarily conducted using publicly available datasets, which could be biased due to variations in how samples were collected and the clinical characteristics involved. Validation in independent cohorts and clinical samples is essential to confirm the clinical relevance of the identified biomarkers. Second, although we identified several promising DEGs, functional validation of these genes in HCC cell lines and animal models is required to better understand their roles in tumorigenesis. Furthermore, while our prognostic model showed promising results, its clinical utility must be evaluated in prospective studies to determine its effectiveness in guiding treatment decisions.

DOI: 10.54254/2753-8818/2025.LD27328

To sum up, this research offers an in-depth examination of the molecular profile of HCC, pinpointing crucial genes and pathways linked to tumor development. The results of our study support the expanding evidence that particular DEGs like EZH2, MCM2, and ALDH2 are prognostic biomarkers in HCC. Creating a prognostic model using Lasso Cox regression improves our comprehension of the disease and may be a useful tool for categorizing patients and tailoring treatments. Further investigations need to concentrate on confirming the roles of these genes and assessing their viability as therapeutic targets in HCC.

References

[1] Bray, F., Ferlay, J., Soerjomataram, I., Siegel, R.L., Torre, L.A. and Jemal, A. (2018) Global Cancer Statistics 2018: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA: A Cancer Journal for Clinicians, 68, 394-424.

[2] Piñero, F., Tanno, M., Aballay Soteras, G., Tisi Baña, M., Dirchwolf, M., Fassio, E., et al. (2020) Argentinian Clinical Practice Guideline for Surveillance, Diagnosis, Staging and Treatment of Hepatocellular Carcinoma. Annals of Hepatology, 19, 546–569.

[3] Feld, J.J. and Krassenburg, L.A.P. (2019) What Comes First: Treatment of Viral Hepatitis or Liver Cancer? Digestive Diseases and Sciences, 64, 1041–1049.

[4] Ringelhan, M., O'Connor, T., Protzer, U. and Heikenwalder, M. (2015) The Direct and Indirect Roles of HBV in Liver Cancer: Prospective Markers for HCC Screening and Potential Therapeutic Targets. Journal of Pathology, 235, 355–367.

[5] Pham, C. and Sin, M.K. (2020) Use of Electronic Health Records at Federally Qualified Health Centers: A Potent Tool to Increase Viral Hepatitis Screening and Address the Climbing Incidence of Liver Cancer. Journal of Cancer Education.

[6] Szabo, G., Saha, B. and Bukong, T.N. (2015) Alcohol and HCV: Implications for Liver Cancer. Advances in Experimental Medicine and Biology, 815, 197–216.

[7] Yao, M., Yang, J.L., Wang, L. and Yao, D.F. (2020) Carcinoembryonic Type Specific Markers and Liver Cancer Immunotherapy. Zhonghua Gan Zang Bing Za Zhi, 28, 466–470.

[8] Henry, S.H., Bosch, F.X. and Bowers, J.C. (2002) Aflatoxin, Hepatitis and Worldwide Liver Cancer Risks. Advances in Experimental Medicine and Biology, 504, 229–233.

[9] Grewal, P. and Viswanathen, V.A. (2012) Liver Cancer and Alcohol. Clinics in Liver Disease, 16, 839–850.

[10] Turati, F., Galeone, C., Rota, M., Pelucchi, C., Negri, E., Bagnardi, V., et al. (2014) Alcohol and Liver Cancer: A Systematic Review and Meta-Analysis of Prospective Studies. Annals of Oncology, 25, 1526–1535.

[11] Saitta, C., Pollicino, T. and Raimondo, G. (2019) Obesity and Liver Cancer. Annals of Hepatology, 18, 810–815.

[12] Lee, Y.C., Cohet, C., Yang, Y.C., Stayner, L., Hashibe, M. and Straif, K. (2009) Meta-Analysis of Epidemiologic Studies on Cigarette Smoking and Liver Cancer. International Journal of Epidemiology, 38, 1497–1511.

[13] Tahmasebi-Birgani, M., Ansari, H. and Carloni, V. (2019) Defective Mitosis-Linked DNA Damage Response and Chromosomal Instability in Liver Cancer. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 1872, 60–65.

[14] Wang, Y., Wang, B., Yan, S., Shen, F., Cao, H., Fan, J., et al. (2016) Type 2 Diabetes and Gender Differences in Liver Cancer by Considering Different Confounding Factors: A Meta-Analysis of Cohort Studies. Annals of Epidemiology, 26, 764–772.

[15] Kudo, M. (2015) Surveillance, Diagnosis, Treatment, and Outcome of Liver Cancer in Japan. Liver Cancer, 4, 39–50.

[16] Yamashita, T. and Kaneko, S. (2016) Liver Cancer. Rinsho Byori, 64, 787–796.

[17] Nozaki, Y., Yamamoto, M., Ikai, I., Yamamoto, Y., Ozaki, N., Fujii, H., et al. (1998) Reconsideration of the Lymph Node Metastasis Pattern (N Factor) from Intrahepatic Cholangiocarcinoma Using the International Union against Cancer TNM Staging System for Primary Liver Carcinoma. Cancer, 83, 1923–1929.

[18] Anwanwan, D., Singh, S.K., Singh, S., Saikam, V. and Singh, R. (2020) Challenges in Liver Cancer and Possible Treatment Approaches. Biochimica et Biophysica Acta (BBA) - Reviews on Cancer, 1873, 188314.

[19] Kaffe, E., Magkrioti, C. and Aidinis, V. (2019) Deregulated Lysophosphatidic Acid Metabolism and Signaling in Liver Cancer. Cancers, 11, 1626.

[20] Mato, J.M., Elortza, F., Lu, S.C., Brun, V., Paradela, A. and Corrales, F.J. (2018) Liver Cancer-Associated Changes to the Proteome: What Deserves Clinical Focus? Expert Review of Proteomics, 15, 749–756.

DOI: 10.54254/2753-8818/2025.LD27328

[21] Wrighton, P.J., Oderberg, I.M. and Goessling, W. (2019) There Is Something Fishy about Liver Cancer: Zebrafish Models of Hepatocellular Carcinoma. Cellular and Molecular Gastroenterology and Hepatology, 8, 347–363.
[22] Charrière, B., Maulat, C., Suc, B. and Muscari, F. (2016) Contribution of Alpha-Fetoprotein in Liver Transplantation for Hepatocellular Carcinoma. World Journal of Hepatology, 8, 881–890.
[23] Lei, Q., Chen, H., Zheng, H., Deng, F., Wang, F., Li, J., et al. (2019) Zygomatic Bone Metastasis from Hepatocellular Carcinoma and the Therapeutic Efficacy of Apatinib: A Case Report and Literature Review. Medicine, 98, e14595.
[24] Szklarczyk, D., Morris, J.H., Cook, H., Kuhn, M., Wyder, S., Simonovic, M., et al. (2017) The STRING Database in 2017: Quality-Controlled Protein-Protein Association Networks, Made Broadly Accessible. Nucleic Acids Research, 45, D362–D368.
[25] Subramanian, A., Kuehn, H., Gould, J., Tamayo, P. and Mesirov, J.P. (2007) GSEA-P: A Desktop Application for Gene Set Enrichment Analysis. Bioinformatics, 23, 3251–3253.
[26] Obuchowski, N.A. and Bullen, J.A. (2018) Receiver Operating Characteristic (ROC) Curves: Review of Methods with Applications in Diagnostic Medicine. Physics in Medicine & Biology, 63, 07TR01.
[27] Tang, C., Jiang, Y., Shao, W., Shi, W., Gao, X., Qin, W., et al. (2016) Abnormal Expression of FOSB Correlates with Tumor Progression and Poor Survival in Patients with Gastric Cancer. International Journal of Oncology, 49, 1489–1496.
[28] Li, H., Li, L., Zheng, H., Yao, X. and Zang, W. (2016) Regulatory Effects of ΔFosB on Proliferation and Apoptosis of MCF-7 Breast Cancer Cells. Tumor Biology, 37, 6053–6063.
[29] Lu, C., Fang, S., Weng, Q., Lv, X., Meng, M., Zhu, J., et al. (2020) Integrated Analysis Reveals Critical Glycolytic Regulators in Hepatocellular Carcinoma. Cell Communication and Signaling, 18, 97.
[30] Dang, C.V. (2012) MYC on the Path to Cancer. Cell, 149, 22–35.
[31] Hsieh, A.L., Walton, Z.E., Altman, B.J., Stine, Z.E. and Dang, C.V. (2015) MYC and Metabolism on the Path to Cancer. Seminars in Cell & Developmental Biology, 43, 11–21.
[32] Stine, Z.E., Walton, Z.E., Altman, B.J., Hsieh, A.L. and Dang, C.V. (2015) MYC, Metabolism, and Cancer. Cancer Discovery, 5, 1024–1039.
[33] Oshi, M. and Takahashi, H. (2020) G2M Cell Cycle Pathway Score as a Prognostic Biomarker of Metastasis in Estrogen Receptor (ER)-Positive Breast Cancer. International Journal of Molecular Sciences, 21, 2921.
[34] De Meyer, T., Bijsmans, I.T., Van de Vijver, K.K., Bekaert, S., Oosting, J., Van Criekinge, W., et al. (2009) E2Fs Mediate a Fundamental Cell-Cycle Deregulation in High-Grade Serous Ovarian Carcinomas. Journal of Pathology, 217, 14–20.