



## OPEN ACCESS

## EDITED BY

Manal S. Fawzy,  
Suez Canal University, Egypt

## REVIEWED BY

Gh Rasool Bhat,  
Virginia Commonwealth University,  
United States  
Xiaonian Zhu,  
Guilin Medical University, China

## \*CORRESPONDENCE

Haitao Zhao,  
✉ Zhaohat@pumch.cn

†These authors have contributed equally to this work

## SPECIALTY SECTION

This article was submitted to Cancer Genetics and Oncogenomics, a section of the journal Frontiers in Genetics

RECEIVED 02 September 2022

ACCEPTED 22 December 2022

PUBLISHED 12 January 2023

## CITATION

Xun Z, Wang Y, Long J, Li Y, Yang X, Sun H and Zhao H (2023), Development and validation of a genomic instability-related lncRNA prognostic model for hepatocellular carcinoma. *Front. Genet.* 13:1034979. doi: 10.3389/fgene.2022.1034979

## COPYRIGHT

© 2023 Xun, Wang, Long, Li, Yang, Sun and Zhao. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

# Development and validation of a genomic instability-related lncRNA prognostic model for hepatocellular carcinoma

Ziyu Xun<sup>†</sup>, Yanyu Wang<sup>†</sup>, Junyu Long<sup>†</sup>, Yiran Li, Xu Yang, Huishan Sun and Haitao Zhao\*

State Key Laboratory of Complex Severe and Rare Disease Department of Liver Surgery, Peking Union Medical College Hospital, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing, China

Genomic instability is a characteristic of tumors, and recent studies have shown that it is related to a poor prognosis of multiple cancers. Long non-coding RNAs (lncRNAs) have become a research hotspot in recent years, and many unknown biological functions are being explored. For example, some lncRNAs play a critical role in the initiation and progression of multiple cancer types by modulating genomic instability. However, the role of genomic instability-related lncRNAs in liver cancer remains unclear. Therefore, we screened genomic instability-related lncRNAs by combining somatic mutation data and RNA-Seq data in The Cancer Genome Atlas (TCGA) database. We established a genomic instability-related lncRNA model (GLncM) involving *ZFPM2-AS1* and *MIR210HG* to predict the hepatocellular carcinoma (HCC) prognosis and further explore the clinical significance of these lncRNAs, and the robustness of the model was validated in the verification set. Thereafter, we calculated the immune score for each patient and explored the relationship between genome instability and the immune microenvironment. The analysis indicated that this model was better than the immune microenvironment in predicting the prognosis of HCC patients, suggesting that the GLncM may be an effective indicator of HCC prognosis and providing a new direction and strategy for estimating the prognosis of HCC patients.

## KEYWORDS

hepatocellular carcinoma, lncRNAs, prognosis, genomic instability, immune environment

## Introduction

Primary liver cancer is the sixth most commonly diagnosed cancer worldwide, and its mortality rate has increased to third (Sung et al., 2021). Primary liver cancer principally includes two histological types: hepatocellular carcinoma (HCC) and intrahepatic cholangiocarcinoma (ICC). Globally, the dominant histological type of primary liver cancer in most countries is HCC, accounting for approximately 75%–85% of all cases of primary liver cancer. In most high-risk areas, including China and Africa, the key risk factors for HCC are chronic hepatitis B virus (HBV) infection and aflatoxin exposure (Sung et al., 2021). In addition to traditional radiotherapy and chemotherapy, targeted therapy and immunotherapy have recently emerged as new treatments for HCC (Kudo et al., 2018; Finn et al., 2020); although these treatments have rapidly progressed, the prognosis of most patients with HCC is still unfavorable due to the heterogeneity of HCC (Allemani et al., 2018). Thus, more sensitive biomarkers must be identified to estimate the prognosis of HCC.

Genomic instability is a prevalent characteristic of most cancers, from precancerous lesions to advanced tumors (Li et al., 2021; Chen et al., 2022). In HCC, HBV-DNA integrates into the host genome at the early stage of tumor amplification and induces genome instability (Levrero and Zucman-Rossi, 2016). Indeed, numerous studies have indicated that genetic instability is related to a poor prognosis of cancers. It has also been reported that the genetically unstable type of breast cancer has a higher risk of significantly shorter recurrence-free survival and metastasis-free survival than other types (Duijf et al., 2019). Consistent with previous reports, genomic instability is also associated with a poor prognosis in colorectal cancer, ovarian cancer and other tumors (Xu et al., 2021; Lakbir et al., 2022). However, studies have shown that extreme genetic instability is correlated with a better prognosis (Andor et al., 2017).

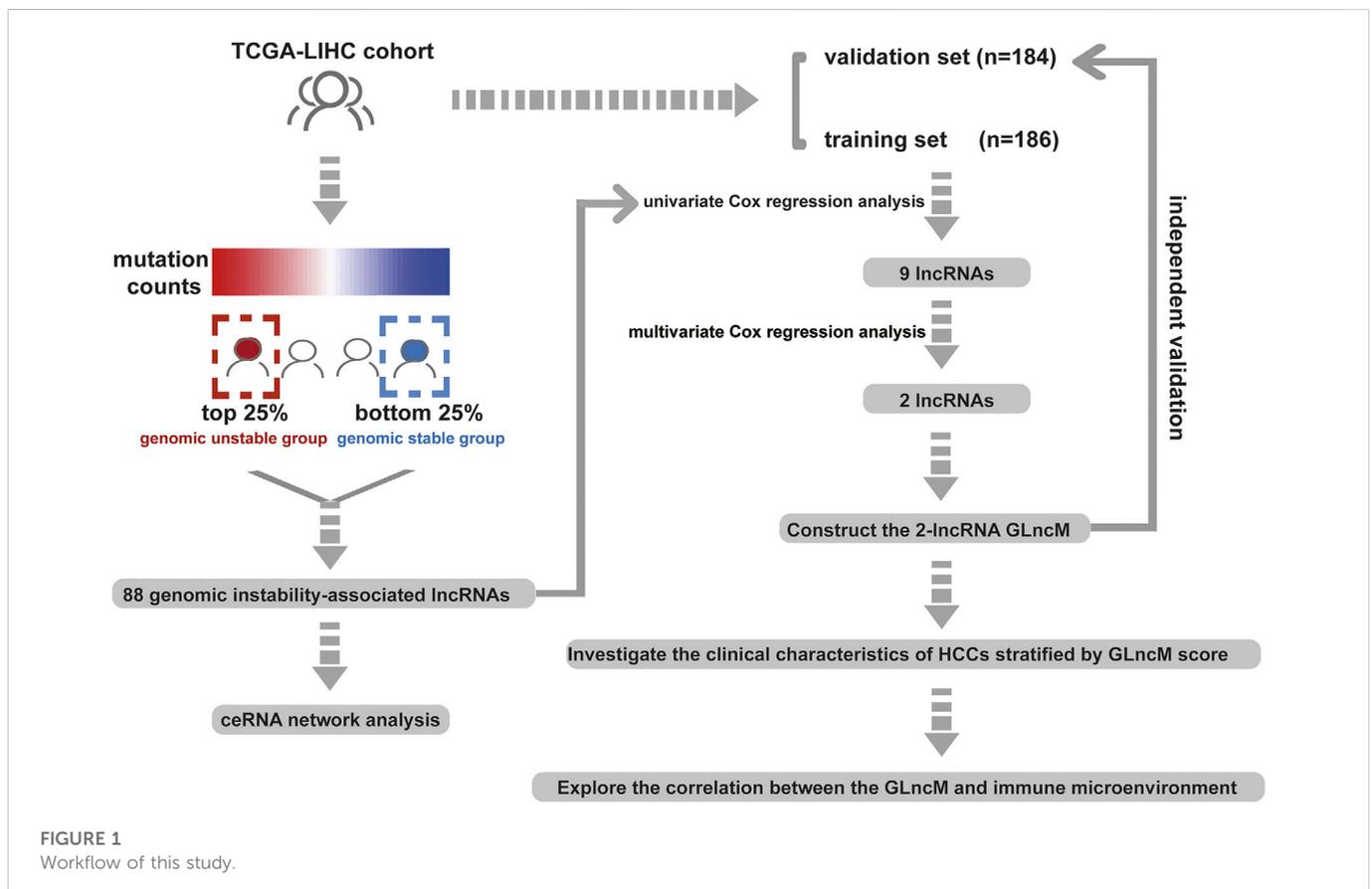
Long non-coding RNAs (lncRNAs) are a class of non-protein-coding RNA molecules whose transcripts are more than 200 nucleotides long (Ulitsky and Bartel, 2013). To date, genome-wide analyses have revealed more than 50,000 genes that transcribe lncRNAs, and their number is still increasing rapidly (Iyer et al., 2015). lncRNAs were initially considered RNAs without biological functions, but researchers have found that some lncRNAs play key roles in tumorigenesis and progression through epigenetic regulation, DNA damage and cell cycle regulation, microRNA regulation, and participation in signal transduction pathways (Hu et al., 2018; Liu et al., 2018; Wang R. et al., 2018; Wang Z. et al., 2018; Huang et al., 2019). For example, H19, HOTAIR and HULC have been shown to facilitate the invasion and metastasis of HCC (Ding et al., 2014; Li et al., 2014; Li D. et al., 2017; Lv et al., 2017; Wang et al., 2017).

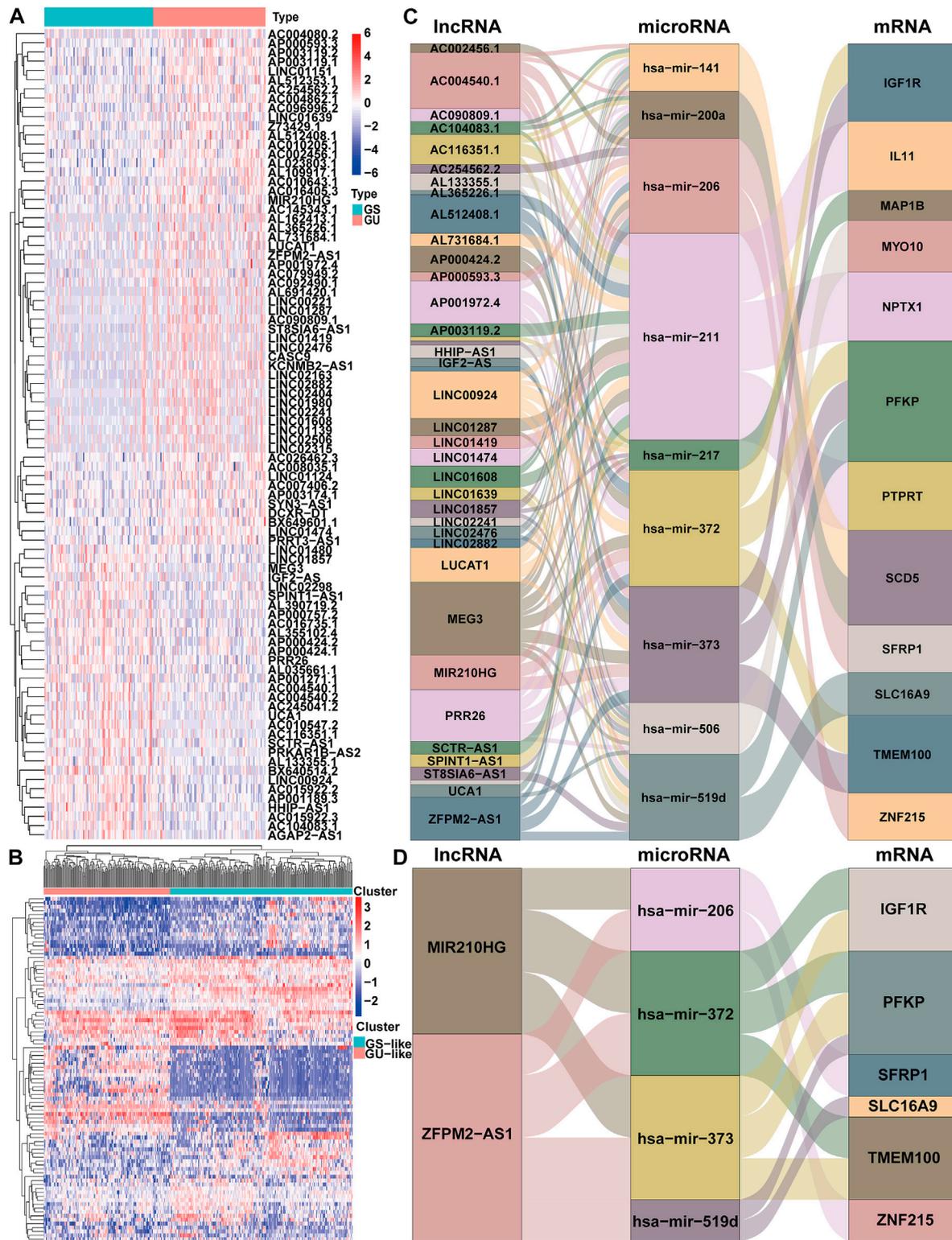
In this article, we combined genetic instability with lncRNAs, classified patients in The Cancer Genome Atlas (TCGA)-LIHC database according to lncRNA expression levels and somatic mutations, and analyzed clinical characteristics to construct an HCC prognostic risk model.

## Results

### Classification of genomic instability-related lncRNAs in HCC patients

The workflow describing the study is shown in Figure 1. We first calculated the total number of mutations in 364 HCC samples and then ranked samples in ascending order to obtain genomic instability-related lncRNAs. The top 25% of the samples ( $n = 90$ ) were selected as the genomic stable group, and the bottom 25% ( $n = 93$ ) were selected as the genomic unstable group. The number of mutations in the genomic stable group ranged from 3 to 78, and the number of mutations in the genomic unstable group ranged from 154 to 2055. Then, we performed a differential expression analysis using the Wilcoxon rank-sum test. As shown in Figure 2A, 88 lncRNAs were differentially expressed between the genomic stable group and the genomic unstable group and were regarded as genomic instability-related lncRNAs. Fifty-six of these lncRNAs were upregulated and 32 lncRNAs were downregulated in the genomic unstable group. We then observed statistically significant differences in the mutation numbers between the genomic stable group and genomic unstable group. In addition,





**FIGURE 2**

Identification and ceRNA network analysis of genomic instability-associated lncRNAs in HCC. **(A)** Differential expression of 88 genomic instability-associated lncRNAs between the genomic stable set and genomic unstable set **(B)** Unsupervised clustering analysis of 364 HCC samples according to the expression patterns of 88 selected lncRNAs correlated with genomic instability. Sankey diagram of the ceRNA network of all lncRNAs correlated with genomic instability **(C)** and Sankey diagram of the ceRNA network of *ZFPM2-AS1* and *MIR210HG* **(D)**. The connection degree of each gene represented by a rectangular box is proportional to the size of the rectangle. GS: genomic stable; GU: genomic unstable.

TABLE 1 Clinical features of the training cohort, testing cohort and whole TCGA-LIHC cohort.

Covariates		TCGA set (n = 370)	Training set (n = 186)	Testing set (n = 184)	p-value
Sex	Female	121 (32.7%)	66 (35.48%)	55 (29.89%)	0.3003
	Male	249 (67.3%)	120 (64.52%)	129 (70.11%)	
Age	≤61	192 (51.89%)	94 (50.54%)	98 (53.26%)	0.6744
	>61	178 (48.11%)	92 (49.46%)	86 (46.74%)	
Family history	No	90 (24.32%)	49 (26.34%)	41 (22.28%)	0.43
	Yes	280 (75.68%)	137 (73.66%)	143 (77.72%)	
Stage	Stage I-II	256 (69.19%)	132 (70.97%)	124 (67.39%)	0.0975
	Stage III-IV	90 (24.32%)	45 (24.19%)	45 (24.46%)	
	Unknown	24 (6.49%)	9 (4.84%)	15 (8.15%)	
Child-pugh classification	A	216 (58.38%)	109 (58.6%)	107 (58.15%)	0.0975
	B-C	22 (5.95%)	11 (5.91%)	11 (5.98%)	
	Unknow	132 (35.68%)	66 (35.48%)	66 (35.87%)	
Grade	G1-G2	232 (62.7%)	114 (61.29%)	118 (64.13%)	0.125
	G3-G4	133 (35.95%)	69 (37.1%)	64 (34.78%)	
	Unknow	5 (1.35%)	3 (1.61%)	2 (1.09%)	
AFP	<400	213 (57.57%)	109 (58.6%)	104 (56.52%)	0.9823
	>400	64 (17.3%)	32 (17.2%)	32 (17.39%)	
	Unknow	93 (25.14%)	45 (24.19%)	48 (26.09%)	

the genomic stable group had higher stromal scores, immune scores, and ESTIMATE scores and better outcomes than the genomic unstable group. Based on the 88 genomic instability-related lncRNAs identified above, 374 HCC samples were grouped into two clusters using an unsupervised hierarchical clustering analysis (Figure 2B). We then observed statistically significant differences in the mutation numbers between the two clusters (131.5 versus 102,  $p < 0.001$ , Wilcoxon rank-sum test, Supplementary Figure S1A). The group with a higher number of mutations was identified as the genomic unstable-like (GU-like) cluster ( $n = 153$ ), and the group with a lower number of mutations was identified as the genomic stable-like (GS-like) cluster ( $n = 221$ ).

We established a ceRNA network based on the lncRNA expression profiles, mRNA expression profiles and microRNA expression data and visualized the network using the ggalluvial R package (Version: 0.12.3) to better understand the function of 88 genomic instability-related lncRNAs (Figure 2C). Seventy genomic instability-related lncRNAs that interact with 86 differentially expressed microRNAs were searched from the miRcode database. Sixteen of these eighty-six differentially expressed microRNAs were obtained from the starBase database. Then, differentially expressed mRNAs were retrieved based on sixteen differentially expressed microRNAs in the miRDB, miRTarBase, and TargetScan databases. According to all three of the databases, twelve differentially expressed mRNAs interacted with nine of the sixteen differentially expressed microRNAs. Finally, 36 genomic instability-related lncRNAs, 9 differentially expressed microRNAs and twelve differentially expressed mRNAs were used to construct a ceRNA network (Figure 2C).

## Construction of a genomic instability-associated lncRNA prognostic model for HCC patients

We constructed a prognostic model based on these lncRNAs to further elucidate the prognostic value of genomic instability-associated lncRNAs. First, 370 HCC samples, excluding duplicate samples, were randomly split into a training set ( $n = 186$ ) and verification set ( $n = 184$ ) at a 1:1 ratio. The clinical features of the training set and verification set are displayed in Table 1. The distribution of each clinical feature between the training set and verification set was balanced. Next, a univariate Cox regression analysis of the training set was conducted to obtain 9 lncRNAs associated with the prognosis ( $p < 0.05$ ; Table 2). Then, a multivariate Cox regression analysis was performed on these 9 lncRNAs to obtain two lncRNAs with independent prognostic value (*ZFPM2-AS1* and *MIR210HG*,  $p < 0.05$ , Table 2). Finally, a genomic instability-related lncRNA model (GLncM) was constructed using the expression levels and coefficients from the multivariate Cox regression analysis of *ZFPM2-AS1* and *MIR210HG* to evaluate the prognosis of HCC patients. The formula was as follows: GLncM score =  $0.119315765508808 \times$  expression level of *ZFPM2-AS1* +  $0.142471464922762 \times$  expression level of *MIR210HG*. In the formula, the coefficients of *ZFPM2-AS1* and *MIR210HG* were both positive, indicating that these two genes were risk factors and that high expression of these genes was correlated with a poorer prognosis than low expression. Liu et al. and Wang et al. previously confirmed that high expression of the two lncRNAs was associated with the proliferation, migration and invasion of HCC cell lines. According

**TABLE 2** Univariate and multivariate Cox regression analysis of the 88 genomic instability-associated lncRNAs correlated with OS in the training cohort.

Gene symbol	Univariate cox regression analysis result					Multivariate cox regression analysis result				
	Coefficient	HR	HR.95L	HR.95H	p-value	Coefficient	HR	HR.95L	HR.95H	p-value
LINC00221	0.138	1.148	1.028	1.282	0.015					
ZFPM2-AS1	0.128	1.136	1.083	1.192	<0.001	0.119	1.127	1.068	1.189	0.000
AC145343.1	0.341	1.407	1.150	1.720	0.001	0.183	1.201	0.953	1.515	0.121
KCNMB2-AS1	0.192	1.212	1.041	1.411	0.013					
PRRT3-AS1	0.093	1.098	1.012	1.191	0.025					
ST8SIA6-AS1	0.087	1.091	1.000	1.190	0.049					
LUCAT1	0.276	1.318	1.129	1.540	<0.001					
MIR210HG	0.197	1.218	1.109	1.338	<0.001	0.142	1.153	1.032	1.289	0.012
CASC9	0.070	1.072	1.029	1.118	0.001					

HR, hazard ratio; OS, overall survival.

to the previous ceRNA network, microRNAs that interact with *ZFPM2-AS1* and *MIR210HG* play a pivotal role in oncogenesis and tumor development (Figure 2D). For instance, Ren et al. previously reported that hsa-mir-206 inhibits tumor growth and metastasis by inhibiting the translation of the SFRP1 protein (Ren et al., 2014). The two lncRNAs in the GLncM may competitively bind to hsa-miR-206, thereby disrupting the binding of the microRNA and SFRP1 and leading to oncogenesis and tumor progression.

Afterward, the risk score of each sample was calculated based on the GLncM, and the median value of the risk score (0.1484198) was regarded as the cutoff value to classify the samples into a high-risk group and a low-risk group. As presented in Figure 3A, the low-risk patients experienced a longer overall survival (OS) than their high-risk counterparts (median OS NA vs. 2.98 years,  $p < 0.001$ ), and the hazard ratio (HR) of the OS of low-versus high-risk group was 0.394 (95% CI: 0.223–0.696). The area under the receiver operating characteristic (ROC) curve (AUC) of the GLncM for OS was 0.742 (95% CI: 0.652–0.833) at 1 year, 0.741 (95% CI: 0.639–0.844) at 3 years and 0.654 (95% CI: 0.52–0.787) at 5 years (Figure 3B). In addition, we ranked the patients according to the risk score in ascending order and displayed the expression patterns of *ZFPM2-AS1* and *MIR210HG* and the landscape of mutation counts and risk scores in the training set (Figure 3C).

## Validation of the GLncM in HCC patients

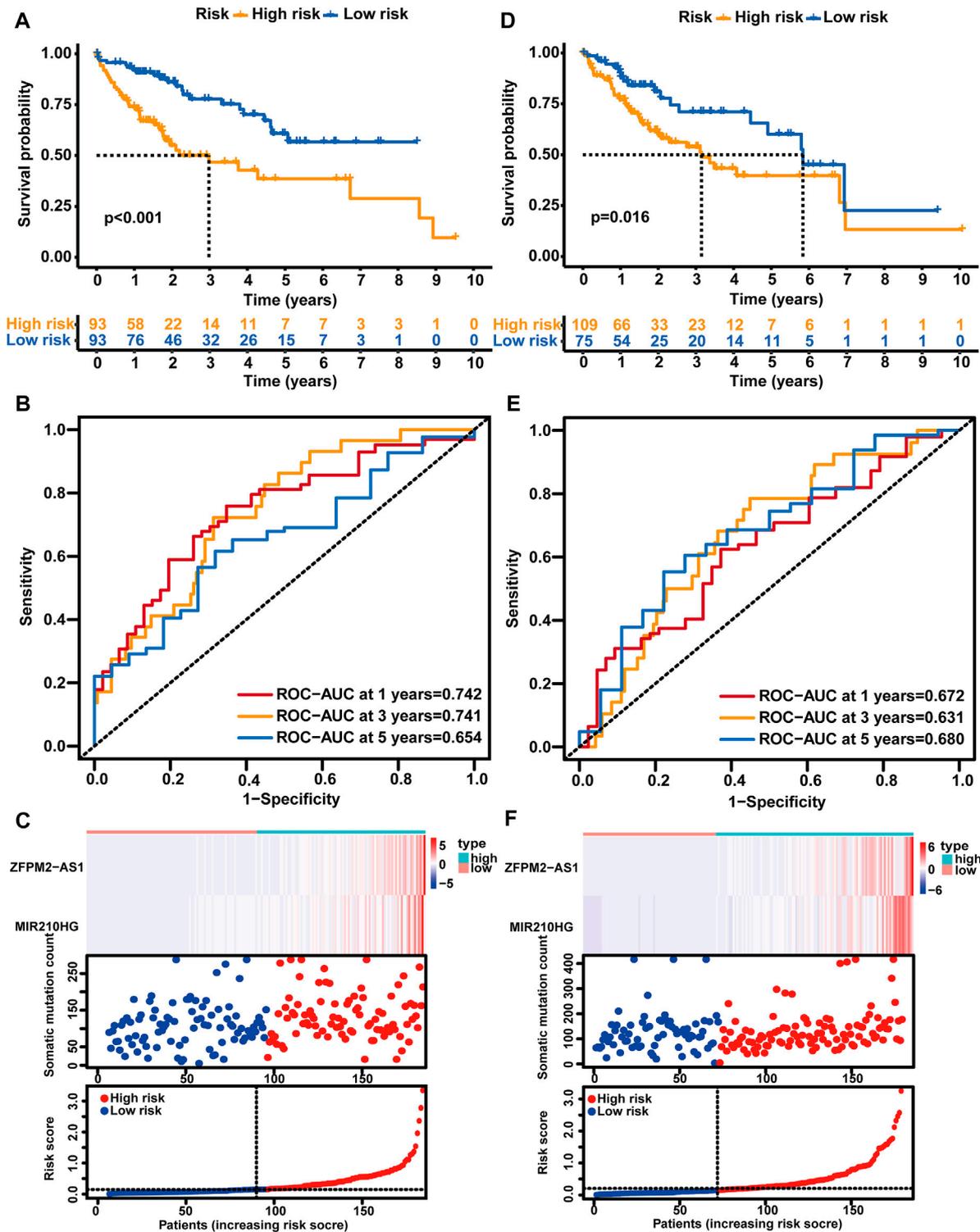
Using the same GLncM formula and cutoff value obtained from the training set, we classified patients in the verification set into high- and low-risk groups. Similar to the result for the training set, the high-risk group experienced a noticeably shorter OS than the low-risk group (median OS 3.15 vs. 5.84 years,  $p = 0.016$ , Figure 3D). The AUC of the GLncM for OS was 0.672 (95% CI: 0.569–0.776) at 1 year, 0.631 (95% CI: 0.515–0.748) at 3 years and 0.680 (95% CI: 0.532–0.829) at 5 years (Figure 3E), and the HR of the OS of the low-versus high-risk group was 0.456 (95% CI: 0.256–0.815). Moreover, we ranked the patients according to the risk score in ascending order and described the *ZFPM2-AS1* and *MIR210HG* expression patterns and the

landscape of mutation counts and risk scores in the testing set (Figure 3F).

Additionally, we stratified all patients in TCGA-LIHC cohort into high- and low-risk patient groups based on the GLncM and the same cutoff value acquired from the training set. Similar to the results described above, patients in the high-risk group experienced a shorter OS than their low-risk counterparts (median OS 3.11 vs. 6.94 years,  $p < 0.001$ , Supplementary Figure S1B). The AUC of the GLncM for OS was 0.709 (95% CI: 0.641–0.778) at 1 year, 0.687 (95% CI: 0.610–0.765) at 3 years and 0.666 (95% CI: 0.568–0.765) at 5 years (Supplementary Figure S1C), and patients in the low-risk group had a lower risk than those in the high-risk group (HR: 0.413, 95% CI: 0.276–0.618). Based on the ascending order of risk scores of patients, we displayed the *ZFPM2-AS1* and *MIR210HG* expression patterns, mutation counts and risk scores for the whole TCGA-LIHC dataset (Supplementary Figure S1D).

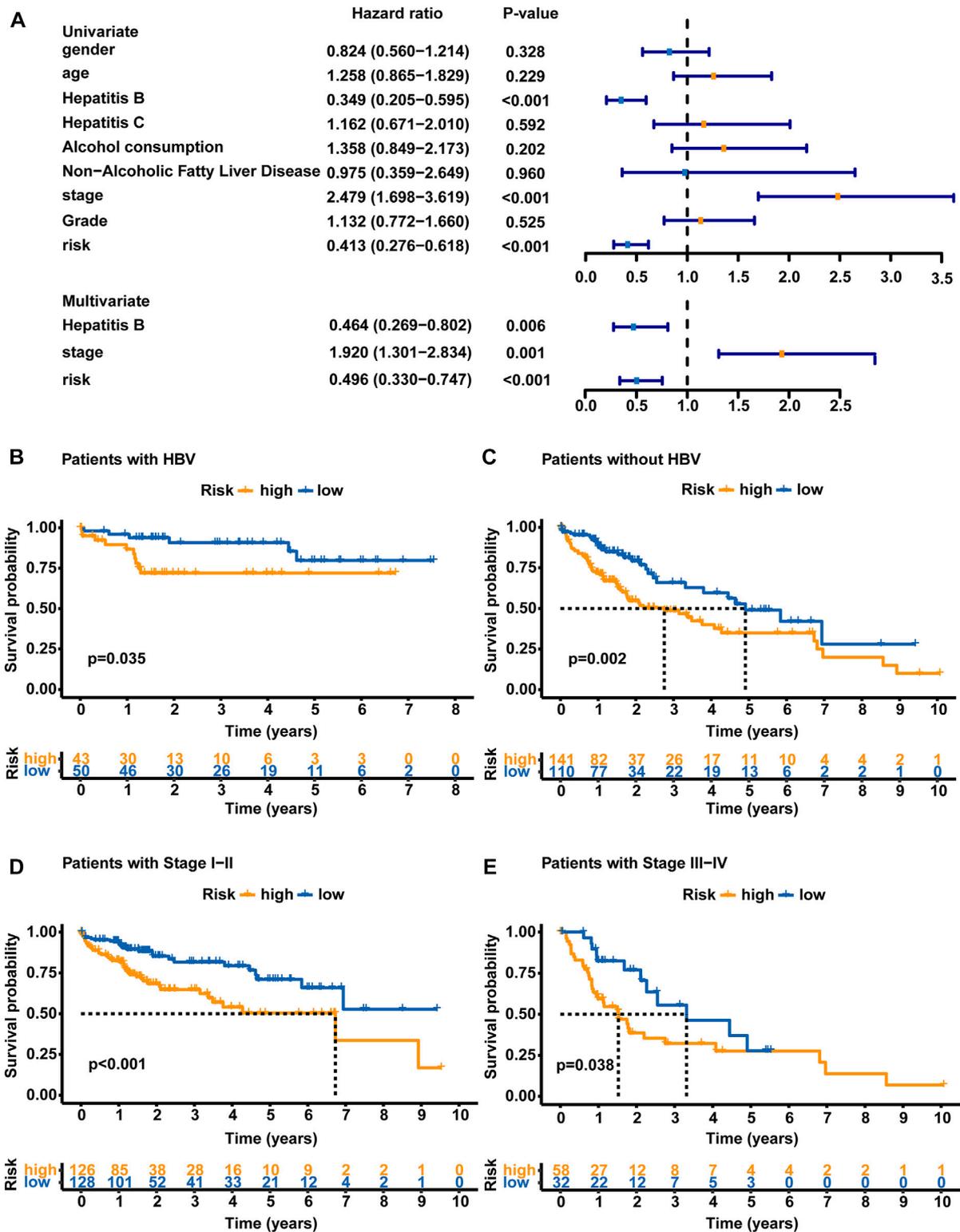
## Independence of the GLncM from other clinical factors

Clinical features, including sex, age, HBV infection status, hepatitis C virus (HCV) infection status, non-alcoholic fatty liver disease, alcohol consumption, pathologic stage and histologic grade, were subjected to a univariate Cox regression analysis and multivariate Cox regression analyses with the established GLncM to assess whether the GLncM was independent of common clinical features. As shown in Figure 4A and Supplementary Table S1, the results of univariate and multivariate Cox regression analyses showed that GLncM was an independent factor for estimating the prognosis of HCC. In addition, the  $p$  values of the other two clinical features, HBV infection status and pathologic stage, were statistically significant in both the univariate and multivariate Cox analyses. Thus, we applied a stratification analysis to verify whether the prognostic value of the GLncM was independent of the HBV infection status and pathologic stage. TCGA-LIHC cohort was classified into the HBV-affected group ( $n = 251$ ) and the non-HBV-affected group ( $n = 93$ ). According to the GLncM and cutoff value identified previously, patients in each HBV infection status group were divided into high- and low-risk subgroups. As



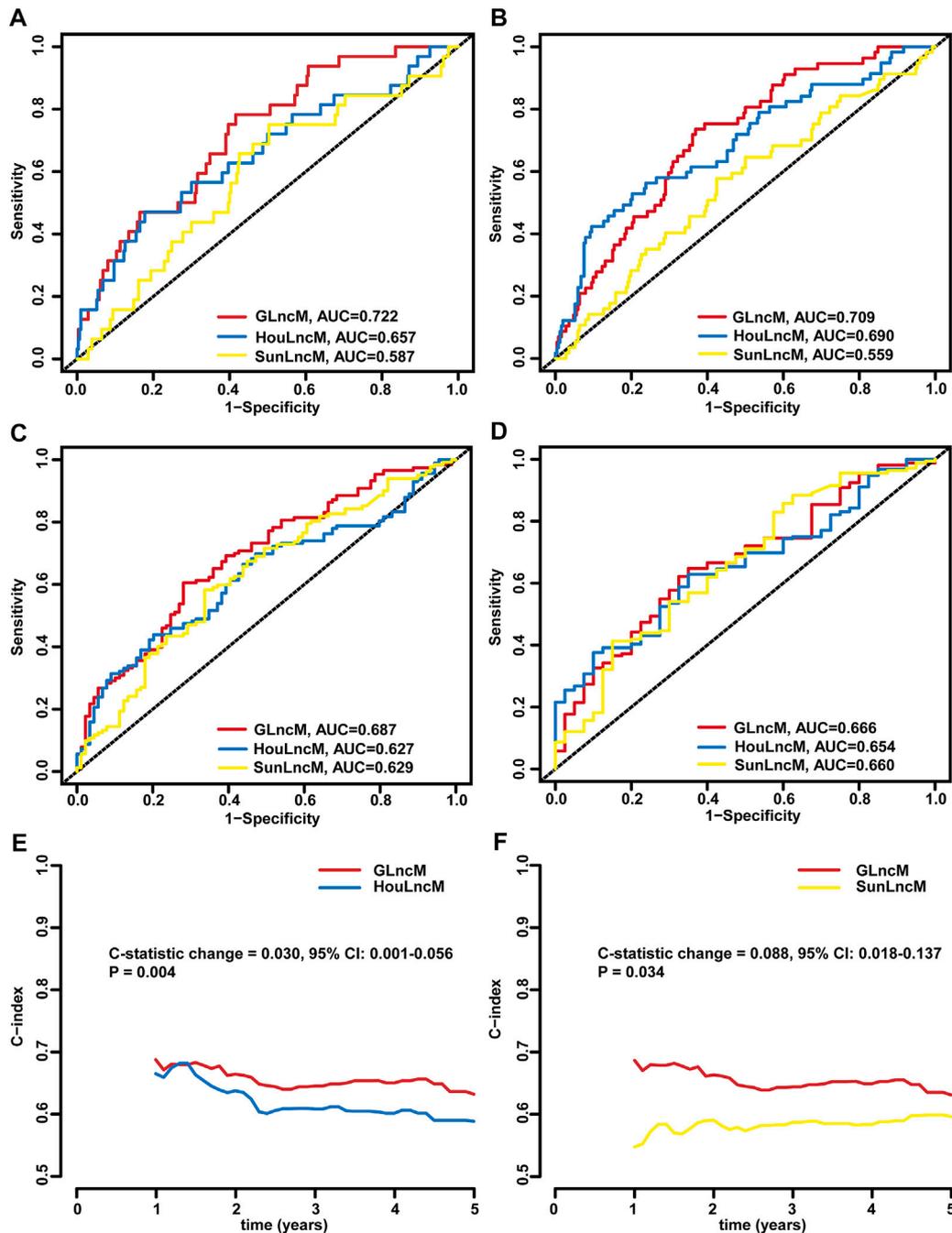
**FIGURE 3**

Analysis of the prognostic performance of the GLncM. Kaplan–Meier survival analysis of the OS of the high-risk and low-risk groups predicted by the GLncM in the training group (A) and verification group (D). The log-rank test was applied for the statistical analysis. (B) and (E) Time-dependent ROC curves for the analysis of the performance of the GLncM at 1, 3, and 5 years in the training group and verification group, respectively. (C) and (F) The lncRNA expression profiles, the risk score and the landscape of mutation counts in the training group and verification group. GLncM: genomic instability-related lncRNA model; OS: overall survival; ROC: receiver operating characteristic.



**FIGURE 4**

Analyses stratified according to the HBV status and tumor stage. (A) Univariate and multivariate Cox regression analyses of the GLncM and clinical characteristics of the whole TCGA-LIHC cohort. Only factors with *p* values less than 0.05 in the univariate analysis were included in the multivariate analysis. OS time illustrated by Kaplan-Meier curves of the high-risk and low-risk subgroups of HCC patients with HBV (B) and without HBV (C). OS time illustrated by Kaplan-Meier curve in high- and low-risk subgroups of stage I-II patients (D) and stage III-IV patients (E). The log-rank test was applied for the statistical analysis. HBV, hepatitis B virus; OS, overall survival.



**FIGURE 5**

Comparison of the performance of GLncM with other models. ROC analysis at 0.5 years (A), 1 year (B), 3 years (C) and 5 years (D) of OS for GLncM, HouLncM and SunLncM. The C-statistic was applied to assess prognostic performance for predicting the OS of HCC patients. The change C-statistic was calculated for GLncM and HouLncM (E) and GLncM and SunLncM (F) in TCGA-LIHC cohort. GLncM: genomic instability-related lncRNA model; HCC: hepatocellular carcinoma; HouLncM: lncRNA-based model from the study by Hou; ROC: receiver operating characteristic; SunLncM: lncRNA-based model from the study by Sun.

described in Figure 4B, a considerable difference in OS was observed between the high- and low-risk subgroups of the HBV-affected group (median OS 2.75 vs. 4.91 years,  $p = 0.035$ ), and similar results were observed in the non-HBV-affected group (median OS NA vs. NA years,  $p = 0.002$ , Figure 4C). Next, patients in TCGA-LIHC dataset were divided according to pathologic stage; patients with stage I or II tumors were grouped into stage I - II ( $n = 254$ ), and those with stage III

or IV tumors were grouped into stage III - IV ( $n = 90$ ). The GLncM with the cutoff value acquired previously divided the patients in the stage I-II group into a high-risk subgroup ( $n = 126$ ) and a low-risk subgroup ( $n = 128$ ), and the high-risk patients experienced a shorter OS than the low-risk patients (median OS 6.73 vs. NA years,  $p < 0.001$ , Figure 4D). Similarly, the GLncM was also applied to divide patients in the stage III - IV group into a high-risk subgroup ( $n = 58$ ) and a low-

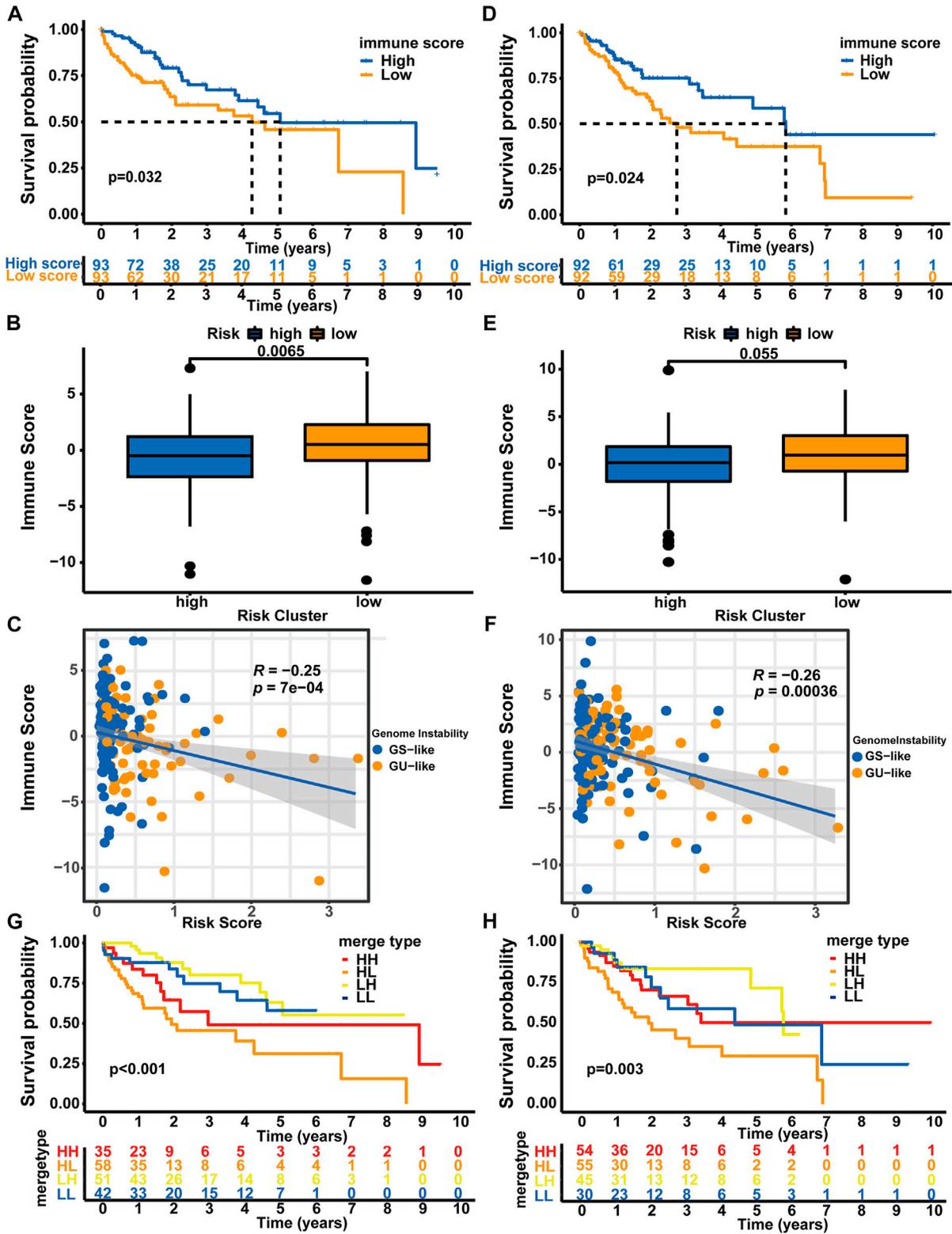


FIGURE 6

Combined analysis of GLncM and the immune microenvironment. Kaplan–Meier curves for the high and low immune score groups in the training set (A) and verification set (D). Difference in the immune scores between the high- and low-risk subgroups in the training set (B) and verification set (E). Scatter plots showing the negative correlation between immune scores and risk scores in the training set (C) and testing set (F). Kaplan–Meier curves for patients in the training set (G) and verification set (H) stratified by both the immune score and risk score. The log-rank test and Spearman’s correlation analysis were applied for statistical analyses. GS-like: genomic stable-like; GU-like: genomic unstable-like; HH: high risk score and high immune score; HL: high risk score and low immune score; LH: low risk score and high immune score; LL: low risk score and low immune score.

risk subgroup ( $n = 32$ ), and a significant difference in OS was observed between the two groups (median OS 1.52 vs. 3.52 years,  $p = 0.021$ , Figure 4E). Based on these results, the GLncM was an independent prognostic factor for HCC.

## Comparison of the GLncM with published lncRNA models in terms of prognostic prediction performance

We further assessed the predictive performance of GLncM by comparing it with two existing lncRNA-related models using the whole TCGA-LIHC dataset: the 5 lncRNA-based model from the study by Hou (HouLncM) (Hou et al., 2018) and the 5 lncRNA-based model from the study by Sun (SunLncM) (Sun et al., 2019).

As shown in Figure 5A, the AUC of GLncM for OS was 0.722 at 0.5 years, which was significantly higher than that of HouLncM (AUC = 0.657) and SunLncM (AUC = 0.587). As shown in Figure 5B, the AUC of GLncM for OS was 0.709 at 1 year, which was significantly higher than that of HouLncM (AUC = 0.690) and SunLncM (AUC = 0.559). The AUC of OS for GLncM was 0.687 at 3 years, which was significantly higher than that of HouLncM (AUC = 0.627) and SunLncM (AUC = 0.629) (Figure 5C). In addition, as shown in Figure 5D, GLncM exhibited higher AUC values (AUC = 0.666) than the other two models (AUC of HouLncM = 0.654, AUC of SunLncM = 0.660) in the whole TCGA-LIHC dataset at 5 years. Moreover, we calculated the C-statistic for GLncM, HouLncM and SunLncM using the CsChange R package, and GLncM improved the C-statistic by 0.030 (95% CI: 0.001–0.056,  $p = 0.004$ ) compared to HouLncM (Figure 5E) and increased the C-statistic by 0.088 (95% CI: 0.018–0.137,  $p = 0.034$ ) compared to SunLncM (Figure 5F).

## The correlation between the immune scores and risk score

Since genomic instability has been reported to play a critical role by creating a precancerous environment (e.g., activation of oncogenic pathways, suppression of oxidative stress responses, and inhibition of immune function) (Rao et al., 2017), we used an immune scoring algorithm similar to that reported by Zhang et al. (2020) to investigate whether the immune environment differed between high-risk patients and low-risk patients. After the immune score of each HCC sample was calculated, the whole TCGA-LIHC cohort was split into a high immune score group and a low immune score group based on the median immune score (0.183). As shown in Supplementary Figure S2A, the OS of the high immune score group was significantly longer than that of the low immune score group (5.84 vs. 3.76 years,  $p < 0.001$ ). Then, we analyzed the distribution of immune scores among the different risk groups. Patients in the high-risk group showed a significantly lower immune score than patients in the low-risk group (median score: 0.114 vs. 0.661,  $p = 0.0017$ , Supplementary Figure S2B). Moreover, using Spearman's correlation analysis, we showed that the immune score was negatively correlated with the risk score (Spearman's correlation coefficient:  $R = -0.25$ ,  $p < 0.001$ , Supplementary Figure S2C).

Afterward, we analyzed the samples from the training set and verification set. Consistent with the results from the whole TCGA-

LIHC dataset, patients in the high immune score group from the training set experienced a longer OS than those in the low immune score group (median OS: 5.07 vs. 4.27 years,  $p = 0.032$ , Figure 6A). Patients in the high-risk group showed a significantly lower immune score than patients in the low-risk group (median score: 0.497 vs. 0.520,  $p = 0.0065$ , Figure 6B), and correlation analyses confirmed that the immune score was negatively associated with the risk score (Spearman's correlation coefficient:  $R = -0.25$ ,  $P = 7e-04$ , Figure 6C). In the testing set, the OS of the low immune score group was significantly shorter than that of the high immune score group (median OS: 2.75 vs. 5.84 years,  $p = 0.024$ , Figure 6D). Patients in the high-risk group showed a significantly lower immune score than patients in the low-risk group (median score: 0.181 vs. 0.973,  $p = 0.055$ , Figure 6E), and correlation analyses confirmed that the immune score was negatively correlated with the risk score (Spearman's coefficient:  $R = -0.26$ ,  $p = 0.00036$ , Figure 6F). We then performed a combined survival analysis according to immune scores and risk scores in the training set, verification set and whole TCGA-LIHC dataset. Figures 6G,H and Supplementary Figure S2D show that patients with high immune scores in the low-risk group experienced a longer OS than other patients. Therefore, the GLncM score has greater prognostic significance than the status of the immune environment.

## Discussion

Currently, the prognosis of patients is mainly determined based on clinical characteristics and biochemical indicators, such as the tumor size, cirrhosis, tumor number, microvascular infiltration and Child–Pugh score (Bertino et al., 2011; Chan et al., 2018; Galle et al., 2019; Huang and Gao, 2020). With the development of emerging detection technologies, researchers have begun to conduct in-depth research on tumors at the molecular level. Specific gene mutations, methylation status, microRNAs, etc., have also been investigated to assess their potential for predicting the prognosis of the disease (Xu et al., 2017; Long et al., 2018; Long et al., 2019). However, the complex relationship between heredity, etiology and environmental risk factors results in genotypic and phenotypic heterogeneity in HCC and has increased the difficulty of determining the prognostic subtypes of HCC (Herath et al., 2006; Feo et al., 2009). Therefore, a more flexible and accurate prognostic model must be developed. With the emergence of high-throughput sequencing technology and microarray analyses, recent studies on lncRNAs have found that they play a role in different mechanisms regulating gene expression. Based on accumulating evidence, lncRNAs are involved in regulating proliferation, apoptosis, invasion and angiogenesis during the occurrence and development of liver cancer (Li et al., 2016; Wang et al., 2016; Zhang et al., 2016). At present, few studies on genomic instability and lncRNAs in liver cancer have been conducted, and the exact biological function and molecular mechanism of related lncRNAs are still unknown. Therefore, further exploration and verification are needed to clarify the complex mechanism.

We performed a combined analysis of gene expression data and somatic mutation data and obtained 88 genomic instability-related lncRNAs. Based on a ceRNA network analysis, we observed that lncRNAs may interrupt the regulatory effect of microRNAs on mRNAs to promote HCC oncogenesis and progression. Several

lncRNAs among these genomic instability-related lncRNAs inhibit metastasis progression. For example, hsa-miR-204 downregulates IL-11 to inhibit the bone metastasis of breast cancer (Pollari et al., 2012).

We further assessed whether lncRNAs correlated with genomic instability predicted the clinical outcomes of HCC patients and acquired a lncRNA model consisting of two lncRNAs correlated with genomic instability (*ZFPM2-AS1* and *MIR210HG*). *ZFPM2-AS1* and *MIR210HG* have previously been reported to be involved in cell proliferation and migration (Li J. et al., 2017; Kang et al., 2019; Wang Y. et al., 2019; He et al., 2020; Liu et al., 2020; Wang et al., 2020), indicating that the model composed of these two lncRNAs is not only an indicator of genomic instability, but also can predict the prognosis in patients with HCC. Based on the results of the ceRNA network analysis, the two lncRNAs included in the GLncM may promote HCC oncogenesis and progression by competing with microRNAs for the binding of mRNAs. For example, the correlation between hsa-miR-372 and TMEM100 plays an important role in the oncogenesis and progression of several types of cancer, including bladder cancer, colorectal cancer, and gastric cancer (Fan and Liu, 2018; Pan et al., 2019; Pang et al., 2019; Wang J. et al., 2019). *ZFPM2-AS1* and *MIR210HG* may disrupt the interaction between hsa-miR-372 and TMEM100 to promote the development of HCC, but the function of hsa-miR-372 and TMEM100 in HCC has not been reported, and thus further experiments are needed to confirm this hypothesis. Furthermore, *ZFPM2-AS1* and *MIR210HG* may also promote the development of HCC by interrupting the interaction between hsa-miR-372 and PFKP, since PFKP has previously been reported to play an important role in cell proliferation (Moon et al., 2011).

Next, we evaluated the synergistic effect of the risk score and immune score on the prognostic stratification of HCC. The results indicated that genomic instability affected the immune microenvironment, but the immune status did not interfere with predictions obtained with the genomic instability-based model.

Currently, reliable prognostic biomarkers for HCC are still lacking. We developed a lncRNA model associated with genomic instability as an independent prognostic marker to stratify HCC patients at risk, which will help clinicians analyze the prognosis of patients using more dimensions and provide new insights into the treatment of patients. Because all samples in this study were collected from databases, future GLncM validation will be conducted in a prospective multicenter cohort study. Our study still has some limitations. Since we did not identify a dataset containing both lncRNAs in the Gene Expression Omnibus (GEO) database, we did not use the external dataset to verify the model. More prospective datasets are needed for the validation and assessment of the performance of the model to ensure its robustness and reproducibility.

## Materials and methods

### Data collection

Data were acquired from TCGA database; these data included somatic mutation data for 364 primary HCC samples, microRNA expression profiles for 375 primary HCC samples, and gene fragments per kilobase of transcript per million reads mapped (FPKM) expression profiles and clinical characteristics for 374 primary HCC samples. In the modeling process, only 370 samples with both mRNA expression data and clinical data were used for

analysis. We annotated the gene symbols and gene types according to the Homo\_sapiens.GRCh38.102. chr.gtf file.

### Classification of genomic instability-related lncRNAs

We used a method similar to that described by Bao et al. to determine the lncRNAs associated with genomic instability (Bao et al., 2020). First, the total number of mutations in each sample was calculated based on somatic mutation data. Then, the samples were sorted in ascending order according to the total number of mutations in each sample. Then, the first 25% of the samples were selected as the genomic stable group, and the last 25% were selected as the genomic unstable group. Then, the Wilcoxon rank-sum test was applied to compare the expression of each lncRNA between the two groups. Finally, genomic instability-related lncRNAs were obtained (false discovery rate (FDR) < 0.05,  $|\log_2\text{fold change (FC)}| > 1$ ).

### Construction of the competing endogenous RNA (ceRNA) network

lncRNA expression data, mRNA expression data and microRNA expression data were used to construct the ceRNA network, and the network was visualized using the ggalluvial R package (Version: 0.12.3) to better illustrate the functions of lncRNAs associated with genomic instability.

The construction of the ceRNA network mainly included two steps. 1) The differentially expressed mRNAs and microRNAs between the genomic stable group and genomic unstable group were obtained using the Wilcoxon rank-sum test and edgeR package (Version: 3.24.3) (Robinson et al., 2010), respectively (FDR < 0.05,  $|\log_2\text{FC}| > 1$ ). 2) The miRcode database (Version: 11) (Jeggari et al., 2012) was applied to determine the interaction between genomic instability-associated lncRNAs and differentially expressed microRNAs. Interactions between differentially expressed microRNAs and differentially expressed mRNAs were determined using the miRDB (Version: 5.0), miRTarBase (Version: 6.1), and TargetScan (Version: 7.2) databases (Agarwal et al., 2015; Wong and Wang, 2015; Chou et al., 2018).

### Construction and validation of a genomic instability-related lncRNA prognostic risk model

First, HCC samples were randomly split into a training set and verification set. Then, in the training set, the prognostic value of genomic instability-related lncRNAs for OS was estimated by performing a univariate Cox regression analysis. The lncRNAs with  $p < 0.05$  were screened in a subsequent multivariate Cox regression analysis. The lncRNAs with  $p < 0.05$  in the multivariate Cox regression analysis were regarded as the final lncRNAs for model construction. Genomic instability-related lncRNAs were obtained by performing univariate and multivariate Cox regression analyses using the survival R package (Version: 3.1–12). The predictive power of the prognostic risk model was evaluated using the log-rank test and Kaplan–Meier survival analysis.

## Immune-related analysis

CIBERSORT (Newman et al., 2015) was used to calculate the infiltration levels of immune cells in each sample. Then, we used a method similar to that described by Zhang et al. (2020) to calculate the immune score for each sample. First, the samples were classified into different immune subtypes based on the CIBERSORT results. Then, differentially expressed genes in the three immune subtypes were obtained using the Wilcoxon rank-sum test. Afterward, the samples were classified into different gene clusters based on these genes. Genes that were positively correlated with gene clusters were regarded as gene signature A; otherwise, they were regarded as gene signature B. The feature genes were identified using the Boruta R package (Version: 7.0.0) (Kursa and Rudnicki, 2010). Principal component analysis (PCA) was conducted for gene signature A and gene signature B of feature genes. The immune score of each sample was calculated by subtracting the score of the first principal component of gene signature A from the score of the first principal component of gene signature B.

## Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga>.

## Author contributions

HZ led the entire project, and all authors participated in the discussion and interpretation of the data and results. ZX, YW, and JL performed bioinformatics analysis, wrote the paper, and were involved in planning the project. XY, YL, and HS participated in the editing of the images and the discussion and interpretation of the data and results.

## References

- Agarwal, V., Bell, G. W., Nam, J. W., and Bartel, D. P. (2015). Predicting effective microRNA target sites in mammalian mRNAs. *Elife* 4, e05005. doi:10.7554/eLife.05005
- Allemani, C., Matsuda, T., Di Carlo, V., Harewood, R., Matz, M., Nikšić, M., et al. (2018). Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet* 391 (10125), 1023–1075. doi:10.1016/s0140-6736(17)33326-3
- Andor, N., Maley, C. C., and Ji, H. P. (2017). Genomic instability in cancer: Teetering on the limit of tolerance. *Cancer Res.* 77 (9), 2179–2185. doi:10.1158/0008-5472.Can-16-1553
- Bao, S., Zhao, H., Yuan, J., Fan, D., Zhang, Z., Su, J., et al. (2020). Computational identification of mutator-derived lncRNA signatures of genome instability for improving the clinical outcome of cancers: a case study in breast cancer. *Brief. Bioinform* 21 (5), 1742–1755. doi:10.1093/bib/bbz118
- Bertino, G., Neri, S., Bruno, C. M., Ardiri, A. M., Calvagno, G. S., Malaguarnera, M., et al. (2011). Diagnostic and prognostic value of alpha-fetoprotein, des-gamma-carboxy prothrombin and squamous cell carcinoma antigen immunoglobulin M complexes in hepatocellular carcinoma. *Minerva Med.* 102 (5), 363–371.
- Chan, A. W. H., Zhong, J., Berhane, S., Toyoda, H., Cucchetti, A., Shi, K., et al. (2018). Development of pre and post-operative models to predict early recurrence of hepatocellular carcinoma after surgical resection. *J. Hepatol.* 69 (6), 1284–1293. doi:10.1016/j.jhep.2018.08.027
- Chen, M., Linstra, R., and van Vugt, M. (2022). Genomic instability, inflammatory signaling and response to cancer immunotherapy. *Biochim. Biophys. Acta Rev. Cancer* 1877 (1), 188661. doi:10.1016/j.bbcan.2021.188661
- Chou, C. H., Shrestha, S., Yang, C. D., Chang, N. W., Lin, Y. L., Liao, K. W., et al. (2018). miRTarBase update 2018: a resource for experimentally validated microRNA-target interactions. *Nucleic Acids Res.* 46 (D1), D296–D302. doi:10.1093/nar/gkx1067
- Ding, C., Cheng, S., Yang, Z., Lv, Z., Xiao, H., Du, C., et al. (2014). Long non-coding RNA HOTAIR promotes cell migration and invasion via down-regulation of RNA binding motif protein 38 in hepatocellular carcinoma cells. *Int. J. Mol. Sci.* 15 (3), 4060–4076. doi:10.3390/ijms15034060
- Duijff, P. H. G., Nanayakkara, D., Nones, K., Srihari, S., Kalimutho, M., and Khanna, K. K. (2019). Mechanisms of genomic instability in breast cancer. *Trends Mol. Med.* 25 (7), 595–611. doi:10.1016/j.molmed.2019.04.004
- Fan, Q., and Liu, B. (2018). Comprehensive analysis of a long noncoding RNA-associated competing endogenous RNA network in colorectal cancer. *Oncotargets Ther.* 11, 2453–2466. doi:10.2147/ott.S158309
- Feo, F., Frau, M., Tomasi, M. L., Brozzetti, S., and Pascale, R. M. (2009). Genetic and epigenetic control of molecular alterations in hepatocellular carcinoma. *Exp. Biol. Med. (Maywood)* 234 (7), 726–736. doi:10.3181/0901-mr-40
- Finn, R. S., Ikeda, M., Zhu, A. X., Sung, M. W., Baron, A. D., Kudo, M., et al. (2020). Phase Ib study of lenvatinib plus pembrolizumab in patients with unresectable hepatocellular carcinoma. *J. Clin. Oncol.* 38 (26), 2960–2970. doi:10.1200/jco.20.00808
- Galle, P. R., Foerster, F., Kudo, M., Chan, S. L., Llovet, J. M., Qin, S., et al. (2019). Biology and significance of alpha-fetoprotein in hepatocellular carcinoma. *Liver Int.* 39 (12), 2214–2229. doi:10.1111/liv.14223
- He, H., Wang, Y., Ye, P., Yi, D., Cheng, Y., Tang, H., et al. (2020). Long noncoding RNA ZFPM2-AS1 acts as a miRNA sponge and promotes cell invasion through regulation of miR-139/GDF10 in hepatocellular carcinoma. *J. Exp. Clin. Cancer Res.* 39 (1), 159. doi:10.1186/s13046-020-01664-1
- Herath, N. I., Leggett, B. A., and MacDonald, G. A. (2006). Review of genetic and epigenetic alterations in hepatocarcinogenesis. *J. Gastroenterol. Hepatol.* 21, 15–21. doi:10.1111/j.1440-1746.2005.04043.x

## Funding

This work was supported by National High Level Hospital Clinical Research Funding (2022-PUMCH-B-128), CAMS Innovation Fund for Medical Sciences (CIFMS) (2022-I2M-C&T-A-003), CAMS Innovation Fund for Medical Sciences (CIFMS) (2021-I2M-1-061) (2021-I2M-1-003), CSCO-hengrui Cancer Research Fund (Y-HR2019-0239) (<http://shenbao.cscoc.org.cn/Review/cscopoInfo/3204> Y-HR2020MS-0415) (<http://shenbao.cscoc.org.cn/Review/cscopoInfo/3203> Y-HR2020QN-0414), CSCO-MSD Cancer Research Fund (Y-MSDZD2021-0213) and National Ten-thousand Talent Program.

## Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

## Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2022.1034979/full#supplementary-material>

- Hou, Y., Yu, Z., Tam, N. L., Huang, S., Sun, C., Wang, R., et al. (2018). Exosome-related lncRNAs as predictors of HCC patient survival: a prognostic model. *Am. J. Transl. Res.* 10 (6), 1648–1662.
- Hu, Y., Lin, J., Fang, H., Fang, J., Li, C., Chen, W., et al. (2018). Targeting the MALAT1/PARP1/LIG3 complex induces DNA damage and apoptosis in multiple myeloma. *Leukemia* 32 (10), 2250–2262. doi:10.1038/s41375-018-0104-2
- Huang, F., and Gao, J. (2020). Modified Child-Pugh grade vs albumin-bilirubin grade for predicting prognosis of hepatocellular carcinoma patients after hepatectomy. *World J. Gastroenterol.* 26 (7), 749–758. doi:10.3748/wjg.v26.i7.749
- Huang, C., Leng, D., Lei, K. C., Sun, S., and Zhang, X. D. (2019). Transcriptome analysis reveals lncRNA-mediated complex regulatory network response to DNA damage in the liver tissue of *Rattus norvegicus*. *J. Cell Physiol.* 234 (12), 23216–23231. doi:10.1002/jcp.28889
- Iyer, M. K., Niknafs, Y. S., Malik, R., Singhal, U., Sahu, A., Hosono, Y., et al. (2015). The landscape of long noncoding RNAs in the human transcriptome. *Nat. Genet.* 47 (3), 199–208. doi:10.1038/ng.3192
- Jeggari, A., Marks, D. S., and Larsson, E. (2012). miRcode: a map of putative microRNA target sites in the long non-coding transcriptome. *Bioinformatics* 28 (15), 2062–2063. doi:10.1093/bioinformatics/bts344
- Kang, X., Kong, F., Huang, K., Li, L., Li, Z., Wang, X., et al. (2019). LncRNA MIR210HG promotes proliferation and invasion of non-small cell lung cancer by upregulating methylation of CACNA2D2 promoter via binding to DNMT1. *Oncol. Targets Ther.* 12, 3779–3790. doi:10.2147/ott.S189468
- Kudo, M., Finn, R. S., Qin, S., Han, K. H., Ikeda, K., Piscaglia, F., et al. (2018). Lenvatinib versus sorafenib in first-line treatment of patients with unresectable hepatocellular carcinoma: a randomised phase 3 non-inferiority trial. *Lancet* 391 (10126), 1163–1173. doi:10.1016/s0140-6736(18)30207-1
- Kursa, M. B., and Rudnicki, W. R. (2010). Feature selection with Boruta package. *J. Stat. Softw.* 36 (11), 1–13. doi:10.18637/jss.v036.i11
- Lakbir, S., Lahoz, S., Cuatrecasas, M., Camps, J., Glas, R. A., Heringa, J., et al. (2022). Tumour break load is a biologically relevant feature of genomic instability with prognostic value in colorectal cancer. *Eur. J. Cancer* 177, 94–102. doi:10.1016/j.ejca.2022.09.034
- Levero, M., and Zucman-Rossi, J. (2016). Mechanisms of HBV-induced hepatocellular carcinoma. *J. Hepatol.* 64, S84–S101. doi:10.1016/j.jhep.2016.02.021
- Li, H., Yu, B., Li, J., Su, L., Yan, M., Zhu, Z., et al. (2014). Overexpression of lncRNA H19 enhances carcinogenesis and metastasis of gastric cancer. *Oncotarget* 5 (8), 2318–2329. doi:10.18632/oncotarget.1913
- Li, T., Xie, J., Shen, C., Cheng, D., Shi, Y., Wu, Z., et al. (2016). Upregulation of long noncoding RNA ZEB1-AS1 promotes tumor metastasis and predicts poor prognosis in hepatocellular carcinoma. *Oncogene* 35 (12), 1575–1584. doi:10.1038/nc.2015.223
- Li, D., Liu, X., Zhou, J., Hu, J., Zhang, D., Liu, J., et al. (2017). Long noncoding RNA HULC modulates the phosphorylation of YB-1 through serving as a scaffold of extracellular signal-regulated kinase and YB-1 to enhance hepatocarcinogenesis. *Hepatology* 65 (5), 1612–1627. doi:10.1002/hep.29010
- Li, J., Wu, Q. M., Wang, X. Q., and Zhang, C. Q. (2017). Long noncoding RNA miR210HG sponges miR-503 to facilitate osteosarcoma cell invasion and metastasis. *DNA Cell Biol.* 36 (12), 1117–1125. doi:10.1089/dna.2017.3888
- Li, H., Zimmerman, S. E., and Weyemi, U. (2021). Genomic instability and metabolism in cancer. *Int. Rev. Cell Mol. Biol.* 364, 241–265. doi:10.1016/bs.irmb.2021.05.004
- Liu, H. T., Liu, S., Liu, L., Ma, R. R., and Gao, P. (2018). EGR1-Mediated transcription of lncRNA-HNF1A-AS1 promotes cell-cycle progression in gastric cancer. *Cancer Res.* 78 (20), 5877–5890. doi:10.1158/0008-5472.Can-18-1011
- Liu, W., Zhang, G. Q., Zhu, D. Y., Wang, L. J., Li, G. T., Xu, J. G., et al. (2020). Long noncoding RNA ZFPM2-AS1 regulates ITGB1 by miR-1226-3p to promote cell proliferation and invasion in hepatocellular carcinoma. *Eur. Rev. Med. Pharmacol. Sci.* 24 (14), 7612–7620. doi:10.26355/eurrev\_202007\_22259
- Long, J., Zhang, L., Wan, X., Lin, J., Bai, Y., Xu, W., et al. (2018). A four-gene-based prognostic model predicts overall survival in patients with hepatocellular carcinoma. *J. Cell Mol. Med.* 22 (12), 5928–5938. doi:10.1111/jcmm.13863
- Long, J., Chen, P., Lin, J., Bai, Y., Yang, X., Bian, J., et al. (2019). DNA methylation-driven genes for constructing diagnostic, prognostic, and recurrence models for hepatocellular carcinoma. *Theranostics* 9 (24), 7251–7267. doi:10.7150/thno.31155
- Lv, M., Zhong, Z., Huang, M., Tian, Q., Jiang, R., and Chen, J. (2017). lncRNA H19 regulates epithelial-mesenchymal transition and metastasis of bladder cancer by miR-29b-3p as competing endogenous RNA. *Biochim. Biophys. Acta Mol. Cell Res.* 1864 (10), 1887–1899. doi:10.1016/j.bbmc.2017.08.001
- Moon, J. S., Kim, H. E., Koh, E., Park, S. H., Jin, W. J., Park, B. W., et al. (2011). Krüppel-like factor 4 (KLF4) activates the transcription of the gene for the platelet isoform of phosphofructokinase (PFKP) in breast cancer. *J. Biol. Chem.* 286 (27), 23808–23816. doi:10.1074/jbc.M111.236737
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337
- Pan, H., Guo, C., Pan, J., Guo, D., Song, S., Zhou, Y., et al. (2019). Construction of a competitive endogenous RNA network and identification of potential regulatory Axis in gastric cancer. *Front. Oncol.* 9, 912. doi:10.3389/fonc.2019.00912
- Pang, X., Zhao, Y., Wang, J., Li, W., Xiang, Q., Zhang, Z., et al. (2019). Competing endogenous RNA and coexpression network analysis for identification of potential biomarkers and therapeutics in association with metastasis risk and progression of prostate cancer. *Oxid. Med. Cell Longev.* 2019, 8265958. doi:10.1155/2019/8265958
- Pollari, S., Leivonen, S. K., Perälä, M., Fey, V., Käkönen, S. M., and Kallioniemi, O. (2012). Identification of microRNAs inhibiting TGF- $\beta$ -induced IL-11 production in bone metastatic breast cancer cells. *PLoS One* 7 (5), e37361. doi:10.1371/journal.pone.0037361
- Rao, C. V., Asch, A. S., and Yamada, H. Y. (2017). Frequently mutated genes/pathways and genomic instability as prevention targets in liver cancer. *Carcinogenesis* 38 (1), 2–11. doi:10.1093/carcin/bgw118
- Ren, J., Huang, H. J., Gong, Y., Yue, S., Tang, L. M., and Cheng, S. Y. (2014). MicroRNA-206 suppresses gastric cancer cell growth and metastasis. *Cell Biosci.* 4, 26. doi:10.1186/2045-3701-4-26
- Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26 (1), 139–140. doi:10.1093/bioinformatics/btp166
- Sun, Y., Zhang, F., Wang, L., Song, X., Jing, J., Zhang, F., et al. (2019). A five lncRNA signature for prognosis prediction in hepatocellular carcinoma. *Mol. Med. Rep.* 19 (6), 5237–5250. doi:10.3892/mmr.2019.10203
- Sung, H., Ferlay, J., Siegel, R. L., Laversanne, M., Soerjomataram, I., Jemal, A., et al. (2021). Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* 71 (3), 209–249. doi:10.3322/caac.21660
- Ulitsky, I., and Bartel, D. P. (2013). lincRNAs: genomics, evolution, and mechanisms. *Cell* 154 (1), 26–46. doi:10.1016/j.cell.2013.06.020
- Wang, X., Zhang, W., Tang, J., Huang, R., Li, J., Xu, D., et al. (2016). LINC01225 promotes occurrence and metastasis of hepatocellular carcinoma in an epidermal growth factor receptor-dependent pathway. *Cell Death Dis.* 7 (3), e2130. doi:10.1038/cddis.2016.26
- Wang, Y., Chen, F., Zhao, M., Yang, Z., Li, J., Zhang, S., et al. (2017). The long noncoding RNA HULC promotes liver cancer by increasing the expression of the HMG2 oncogene via sequestration of the microRNA-186. *J. Biol. Chem.* 292 (37), 15395–15407. doi:10.1074/jbc.M117.783738
- Wang, R., Ma, Z., Feng, L., Yang, Y., Tan, C., Shi, Q., et al. (2018). LncRNA MIR31HG targets HIF1A and P21 to facilitate head and neck cancer cell proliferation and tumorigenesis by promoting cell-cycle progression. *Mol. Cancer* 17 (1), 162. doi:10.1186/s12943-018-0916-8
- Wang, Z., Yang, B., Zhang, M., Guo, W., Wu, Z., Wang, Y., et al. (2018). lncRNA epigenetic landscape analysis identifies EPIC1 as an oncogenic lncRNA that interacts with MYC and promotes cell-cycle progression in cancer. *Cancer Cell* 33 (4), 706706–706720. doi:10.1016/j.ccell.2018.03.006
- Wang, J., Zhang, C., Wu, Y., He, W., and Gou, X. (2019). Identification and analysis of long non-coding RNA related miRNA sponge regulatory network in bladder urothelial carcinoma. *Cancer Cell Int.* 19, 327. doi:10.1186/s12935-019-1052-2
- Wang, Y., Li, W., Chen, X., Li, Y., Wen, P., and Xu, F. (2019). MIR210HG predicts poor prognosis and functions as an oncogenic lncRNA in hepatocellular carcinoma. *Biomed. Pharmacother.* 111, 1297–1301. doi:10.1016/j.biopha.2018.12.134
- Wang, A. H., Jin, C. H., Cui, G. Y., Li, H. Y., Wang, Y., Yu, J. J., et al. (2020). MIR210HG promotes cell proliferation and invasion by regulating miR-503-5p/traf4 axis in cervical cancer. *Aging (Albany NY)* 12 (4), 3205–3217. doi:10.18632/aging.102799
- Wong, N., and Wang, X. (2015). miRDB: an online resource for microRNA target prediction and functional annotations. *Nucleic Acids Res.* 43, D146–D152. doi:10.1093/nar/gku1104
- Xu, R. H., Wei, W., Krawczyk, M., Wang, W., Luo, H., Flagg, K., et al. (2017). Circulating tumour DNA methylation markers for diagnosis and prognosis of hepatocellular carcinoma. *Nat. Mater* 16 (11), 1155–1161. doi:10.1038/nmat4997
- Xu, F., Liu, T., Zhou, Z., Zou, C., and Xu, S. (2021). Comprehensive analyses identify APOBEC3A as a genomic instability-associated immune prognostic biomarker in ovarian cancer. *Front. Immunol.* 12, 749369. doi:10.3389/fimmu.2021.749369
- Zhang, J. Y., Weng, M. Z., Song, F. B., Xu, Y. G., Liu, Q., Wu, J. Y., et al. (2016). Long noncoding RNA AFAP1-AS1 indicates a poor prognosis of hepatocellular carcinoma and promotes cell proliferation and invasion via upregulation of the RhoA/Rac2 signaling. *Int. J. Oncol.* 48 (4), 1590–1598. doi:10.3892/ijo.2016.3385
- Zhang, X., Shi, M., Chen, T., and Zhang, B. (2020). Characterization of the immune cell infiltration landscape in head and neck squamous cell carcinoma to aid immunotherapy. *Mol. Ther. Nucleic Acids* 22, 298–309. doi:10.1016/j.omtn.2020.08.030