

Comprehensive analysis of five long noncoding RNAs expression as competing endogenous RNAs in regulating hepatoma carcinoma

Xin Lou | Jun Li | Dong Yu | Ya-Qing Wei | Shuang Feng | Jin-Jin Sun 

Department of Hepatopancreatobiliary Surgery, Tianjin Medical University Second Hospital, Tianjin Medical University, Tianjin, China

Correspondence

Jin-Jin Sun, Department of Hepatopancreatobiliary Surgery, Tianjin Medical University Second Hospital, Tianjin Medical University, 23 Pingjiang road, Hexi district, Tianjin, China.
Email: jsun02@tmu.edu.cn

Funding information

Natural Science Foundation of Tianjin City, Grant/Award Number: 16JCYBJC27100

Abstract

Liver cancer is the most common cancer and is the epitome of a recalcitrant cancer. Increasing evidence shown that long noncoding RNAs (lncRNA) were associated with cancer-related death and could function as a competing endogenous RNA (ceRNA). To explore regulatory roles and potential prognostic biomarkers of lncRNA for liver cancer, RNA-sequencing expression data were downloaded from the TCGA database and GEO database. A total of 357 patients were randomly divided into a discovery group and a validation group, of which 313 patients can obtain clinical data. In discovery phrase, 58 lncRNAs, 16 miRNAs, and 34 mRNAs were screened to construct the ceRNA network based on 252 patients employed from discovery group. Univariate and multivariate Cox hazard regression analysis model revealed that five lncRNAs (AATK-AS1, C10orf91, LINC00162, LINC00200, and LINC00501) from 58 lncRNAs were formulated to predict the overall survival (OS). We used the value of gene expression and regression coefficients to construct a risk score based on the five lncRNAs. Next, we validated our model in the GSE116174 dataset ($n = 64$) and the validation group ($n = 94$) from TCGA database. Subgroup analysis suggest that the five lncRNAs played critical parts in early stage in cancer from both discovery and validation groups. The five lncRNAs were also found to be associated with immune cells infiltration including CD4⁺ memory activated, NK cells activated and mast cells activated, then the results were also validated according to the validation group. Furthermore, KEGG pathway enrichment analysis showed that these nine coexpressed modules using the method of WGCNA, and many of these pathways are associated with the development and progression of disease. At last, the transcription factor binding sites (TFBS) of the five lncRNAs were predicted, which help us to understand the potential mechanism that the TFBS adjusted the ceRNA network. In summary, the ceRNA regulatory network may contribute to a better understanding of liver cancer mechanism and provide potential prognostic biomarkers and therapeutic targets.

KEYWORDS

ceRNA, hepatoma carcinoma, lncRNA, prognosis, TCGA

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2019 The Authors. *Cancer Medicine* published by John Wiley & Sons Ltd.

1 | INTRODUCTION

Hepatocellular carcinoma (HCC) became the sixth common cancer and the death rate increased year by year.¹ And the rate accounts for a higher proportion in developing countries including China due to the high prevalence of chronic hepatitis C.² There currently exists a lot of ways to treatment including surgical resection, transplantation, and local ablation for early liver cancer,³ but the overall survival did not present apparently variation. With the development of molecular multi-kinase inhibitors, sorafenib, regorafenib, and lenvatinib have increased the overall survival (OS) rate of HCC and have been approved by the US Food and Drug Administration (FDA) to cure HCC.⁴ However, those drugs can only improve less than 4 months OS in advanced HCC and did not response a better prognosis.^{4,5} Therefore, the development of novel treatment, the identification of new prognostic biomarkers and a clearer understanding of molecular mechanisms are essential and urgently required.

Noncoding RNA (ncRNA) sequences include small nucleolar RNAs, long noncoding RNA (lncRNA), miRNA, and small interfering RNA (siRNA), of which lncRNAs are >200 nucleotides in length and regulate gene expression at the levels of chromatin organizational, transcriptional, or posttranscriptional.⁶ miRNAs-gene-regulatory ncRNA could direct RNA-induced silencing complex (RISC) miRNA response elements (MRE), which repressed protein production through inhibiting translation or destabilizing the mRNA.⁷ MRE located in 3' untranslated region (UTR), coding sequence (CDS), and 5'UTR, and could be found on lncRNA and mRNA. As known, each miRNA has various RNA targets, which has led to the hypothesis that the different RNAs sharing the same MRE compete with each other for limited miRNA,^{8,9} so that acting as competitive endogenous RNAs (ceRNA) and regulating gene expression. lncRNAs are extensively targeted by miRNAs through 3' UTR or 5'UTR, meaning that they could serve as ceRNAs,¹⁰ the lncRNA as ceRNA were associated with cellular biological process and also serve important roles in tumorigenesis.¹¹ Recent published studies confirmed the ceRNA theory involved in the progression of various types of cancer.¹²⁻¹⁵ These studies aimed at some genes including mRNA, miRNA, and lncRNA associated with the cancer and confirmed that some of genes serve important roles in treatment and prognosis. With the development of experimental studies and techniques for lncRNA discovery, lncRNA-associated ceRNA networks have been constructed and analyzed in colorectal cancer, gastric cancer, and osteosarcoma.^{13,16,17} However, there existed fewer analysis of ceRNA network in hepatocellular carcinoma (HCC).

In the present study, we conducted a comprehensive analysis of mRNA, lncRNA, and miRNA expression profiles in HCC and the lncRNA-sequencing (lncRNA-seq) data, mRNA-seq, and miRNA-seq expression of HCC samples were downloaded from TCGA database and the differently

expressed RNAs were screened to construct ceRNA network. Furthermore, univariate and multivariate Cox regression analysis was further conducted to establish a risk assessment system based on the regression coefficient. Subsequently, the assessment model was validated in validation and entire group, and we further explore biological function as well as immune cells infiltration characters associated with the five lncRNAs, then we predicted the TFBS of HCC which may regulate the ceRNA network to understand the potential mechanism.

2 | MATERIAL AND METHOD

2.1 | Patient information and preprocessing

The data of RNA-seq expression and clinical information were downloaded from the TCGA database (<https://portal.gdc.cancer.gov/>) and GEO database (GSE116174). The patients obtained from TCGA were randomly divided into a discovery group and a validation group. The data downloaded from GEO were used as a validation group. The discovery group was used to construct model, and the validation group was used to validate the efficiency of the model. Firstly, we obtained the lncRNA expression based on annotation of Genecode (<https://www.gencodegenes.org/>) by screening them from the mRNA expression profiles we have downloaded. Consequently, the RNA-sequencing data of TCGA covered 19767 mRNA, 14718 lncRNA, and 1881 miRNA. Next, after we conducted normalization of RNA-seq of TCGA data and GEO data, the differentially expressed mRNAs (DEmiRNA), lncRNAs (DElncRNA), and miRNA (DEmiRNA) were conducted based on Bioconductor package of edgeR in R 3.5.2 with the threshold of \log_2 fold change > 2 and $P < .01$.

2.2 | Construction of ceRNA network

The miRcode was used to predict the interaction of DElncRNA with DEmiRNA, and the mRNAs were retrieved according to miRTarBase, TargetScan, and miRDB based on targeted miRNA. To increase the reliability of the results, only miRNA-mRNA interaction found in all three databases were selected as candidate genes for constructing the ceRNA network. Then, the obtained mRNA intersected with the DE mRNAs to screen final targeted mRNAs. Next, the lncRNA-miRNA-mRNA ceRNA network was constructed. At last, the interactions and visualization were conducted by the Cytoscape software (<https://cytoscape.org/>).

2.3 | Risk assessment model construction and evaluation

After ceRNA network was constructed, we obtained lncRNA to univariable Cox regression analysis to select lncRNA

associated with OS of patients with liver cancer, only the lncRNA with a statistical significance (P value $<.05$) were enrolled into multivariable Cox regression. A patient risk assessment model was constructed through the regression coefficients with lncRNA expression. In other words, the risk score was the linear combination of expression value of selected lncRNAs weighted by the regression coefficients. A risk score of patients was calculated based on the Equation. Risk score = $\text{Exp1} * \text{Coe1} + \text{Exp2} * \text{Coe2} + \text{Exp3} + \text{Coe3} + \dots + \text{Exp}_i * \text{Coe}_i$.¹⁸ In this Equation, the Exp are the expression value of lncRNAs and Coe are their corresponding coefficients from the multivariable Cox regression analysis.

2.4 | Prognostic survival analysis

The risk scores of HCC patients were calculated according to above risk assessment system, the patients were divided into high risk and low risk using the median risk score as boundary. The Kaplan-Meier method was used to assess the efficiency on OS in high-risk and low-risk patients. The P value of log-rank test less than .05 was considered as significance. We conducted the prognostic survival analysis on all this three groups including discovery group, validation group from TCGA, validation group from GEO. Then, we performed an entire analysis to combine the all sample from these three groups.

2.5 | Pathway enrichment analysis according to weighted correlation network analysis (WGCNA)

We conducted a coexpression network using Bioconductor package of “WGCNA” in R 3.5.2 to find the gene modules closed to our risk scores. The thresholding power was selected to 5 and the genes were clustered into nine modules based on clinical characters. The most significant modules associated with risk score were selected and the genes enrolled into this module were used to conduct biology processes analysis and pathway enrichment analysis. The biology processes analysis was performed using the online web tool “DAVID” (<https://david.ncifcrf.gov/>). The pathway enrichment analysis was performed according to the Bioconductor package of “clusterProfiler” in R 3.5.2.

2.6 | Evaluation of tumor infiltrating immune cells and the relation of immune cells with five lncRNAs

To infer the infiltrating immune cells associated with five lncRNAs, we used the targeted mRNAs to predict the proportion of 22 types of infiltrating immune cells using the CIBERSORT web portal (<https://cibersort.stanford.edu/index.php>) which is a gene expression-based deconvolution

algorithm.¹⁹ We can obtain the significant immune cell type and the difference in immune cells between cancer and normal tissue, and we further study the relation of these immune cells with our model to evaluate the OS in patients with HCC.

2.7 | Transcription regulation prediction on ceRNA network

To further understand the mechanisms of ceRNA network in HCC, we predict transcription factor binding sites (TFBS) which regulate the ceRNA network in HCC. We search the promoter region of gene on the basis of the web tool “NCBI.” At first, the potential promoter region is generally thought the region from the sequence of 2000 bp upstream to 100 bp downstream of the starting gene point. Then, the TFBS would be predicted using web tool “UCSC” and “JASPAR.” The intersections of TFBS among the ceRNA network including lncRNA associated with prognosis were calculated respectively, the TFBS were thought to regulate ceRNA network. At last, we conducted KEGG pathways analysis based on these TFBS, which help us to better understand the mechanism of the ceRNA with TFBS.

3 | RESULTS

3.1 | Data source and identification of DERNAs

A total of 357 patients with liver cancer and 14748 lncRNA expression values were collected from TCGA database. Patients were randomly divided into a discovery group ($n = 252$) and validation group ($n = 105$). The discovery group including 33 normal and 219 tumor tissue patients obtained 1035 DElncRNAs according to the criteria ($P < .01$ and $\log_2\text{FCI} > 2$). The volcano plot is presented in Figure 1. A total of 123 DEmiRNAs and 1986 DEMRNAs were obtained from TCGA database. And the results are shown in Figure 1.

3.2 | Construction of lncRNA-miRNA-mRNA ceRNA network

We assessed the relationship between miRNAs and DElncRNAs on basis of miRcode downloaded from the website (<http://www.mircode.org/>) which present correspondence between lncRNAs and miRNAs. The target mRNA of DEmiRNA was predicted according to the intersection of these three databases (TargetScan, miRDB, and miRarBase). At last, 58 lncRNAs, 16 miRNAs, and 34 mRNAs were included to construct ceRNA network, and the visualization of coexpression was built using the software of Cytoscape (Figure 2).

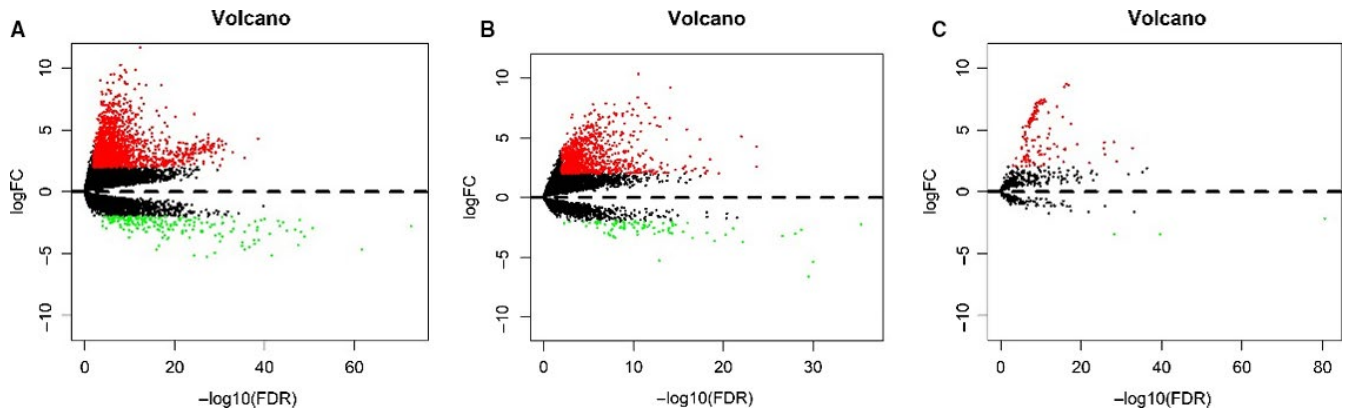


FIGURE 1 A Volcano plot of differentially expressed RNAs in patients with hepatocellular carcinoma ($\log_2FCI > 2$ and $P < .01$). A, DE mRNAs; B, DE lncRNAs; C, DE miRNA. up-regulated RNA and down-regulated was represented in red dot and green dot respectively

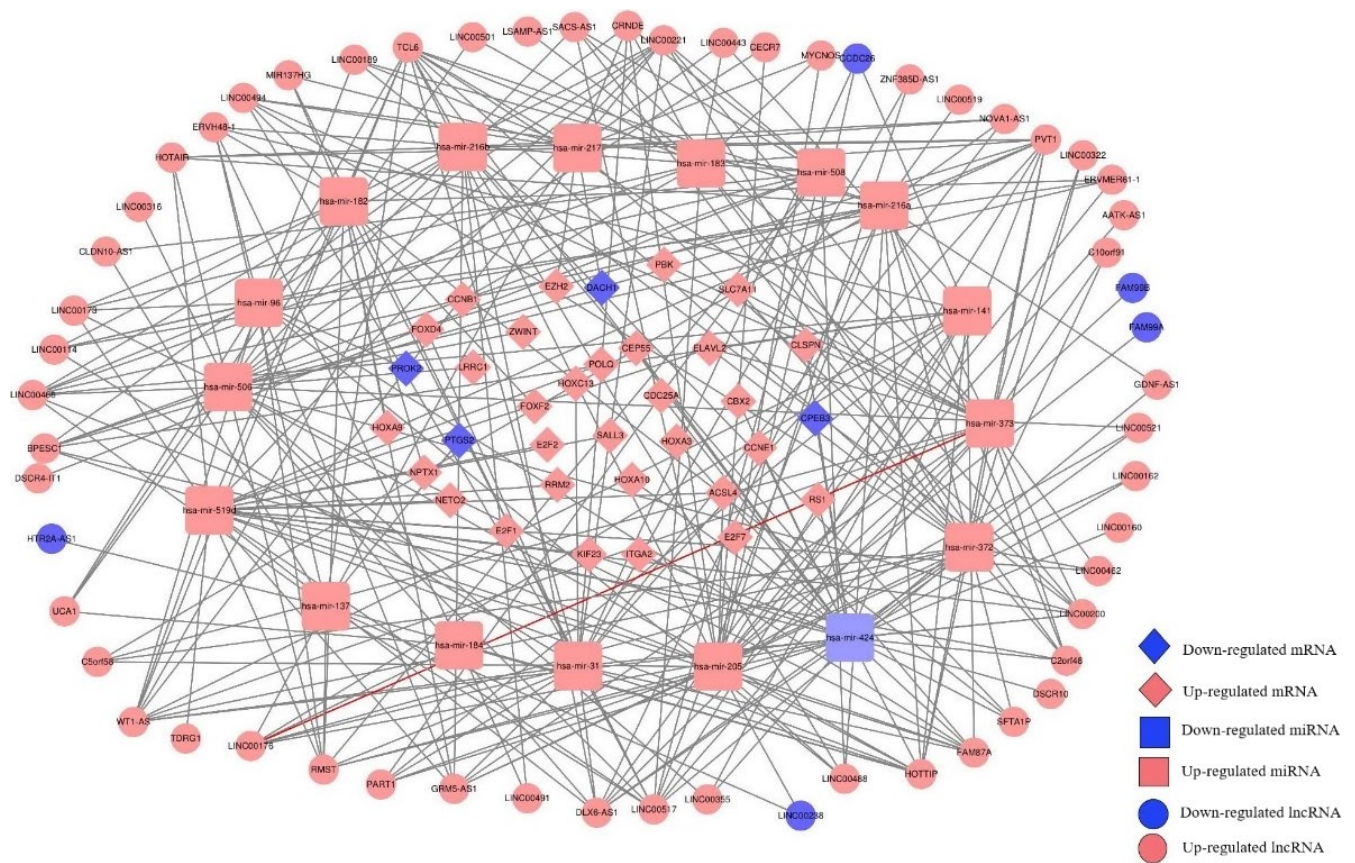


FIGURE 2 The visualization of ceRNA network in hepatoma carcinoma. The lncRNA-miRNA-mRNA ceRNA network

3.3 | Screening for lncRNAs as biomarkers related to overall survival and prognosis

Based on the ceRNA network, a total of 58 differentially expressed lncRNA were analyzed by the Univariate and Cox hazards regression analyses. Twenty-four lncRNAs were

identified to be significantly correlated with prognosis based on univariate hazards regression analysis (P value $< .05$), of which these lncRNAs were screened to conduct multivariate Cox regression analysis. We finally obtained five lncRNAs, namely, AATK-AS1, C10orf91, LINC00162, LINC00200, and LINC00501 (Table 1). On the basis of

TABLE 1 Five lncRNAs significantly associated with the overall survival in patients with liver cancer in the discovery group

Gene name	Ensembl ID	Univariate analysis		Multivariate analysis	
		HR	<i>P</i> values	HR	<i>P</i> values
AATK-AS1	ENSG00000225180	1.153	.021	1.178	.016
C10orf91	ENSG00000180066	1.247	<.001	1.181	<.001
LINC00162	ENSG00000275874	1.266	<.001	1.278	<.001
LINC00200	ENSG00000229205	1.246	<.001	1.271	<.001
LINC00501	ENSG00000203645	1.191	.027	1.198	.021

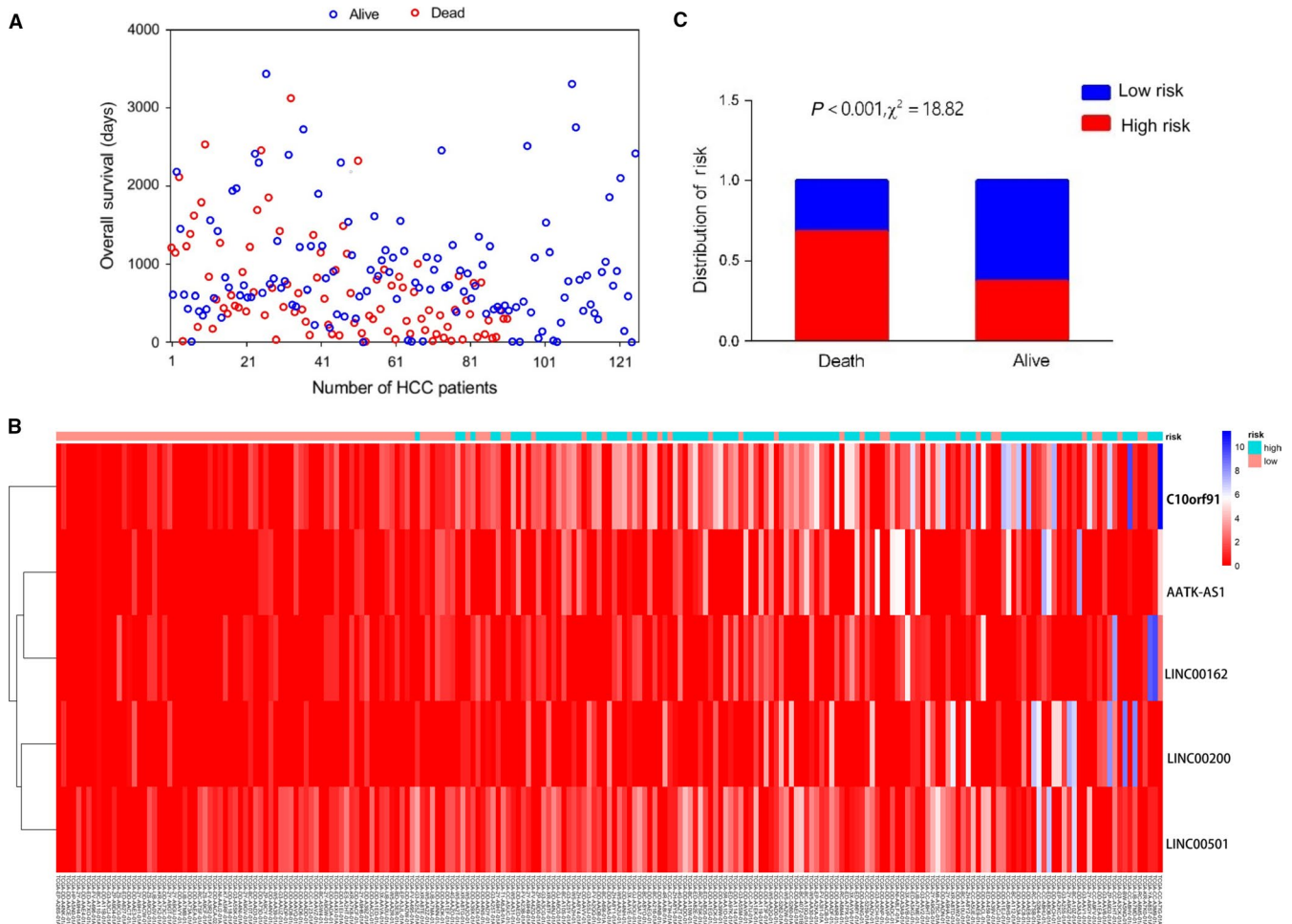


FIGURE 3 The five-lncRNA model and its prognostic value for liver cancer. A, survival time of patients with different vital status. B, Heatmap of five lncRNAs expression between low-risk score and high-risk score. C, The risk attribution in Death group and Alive group

multivariable Cox, a risk score was constructed as following: risk score = $1.178 \times \exp(\text{AATK-AS1}) + 1.181 \times \exp(\text{C10orf91}) + 1.278 \times \exp(\text{LINC00162}) + 1.271 \times \exp(\text{LINC00200}) + 1.198 \times \exp(\text{LINC00501})$. The heatmap revealed that the expression level of five lncRNAs varied as the risk scores (Figure 3). Our data showed that mortality rate in high-risk group was significantly higher than low-risk group (Figure 3), which indicates the five lncRNAs play a critical role in liver cancer.

3.4 | The prognostic values of five lncRNAs in discovery and validation group

Our data in discovery group showed that the patients who had low-risk scores present a longer OS time than higher risk group (Figure 4A). To validate above finding, we employed the validation group from TCGA and a GSE116174 dataset. The two validation groups were consistent with the result of discovery group. The entire samples were employed together

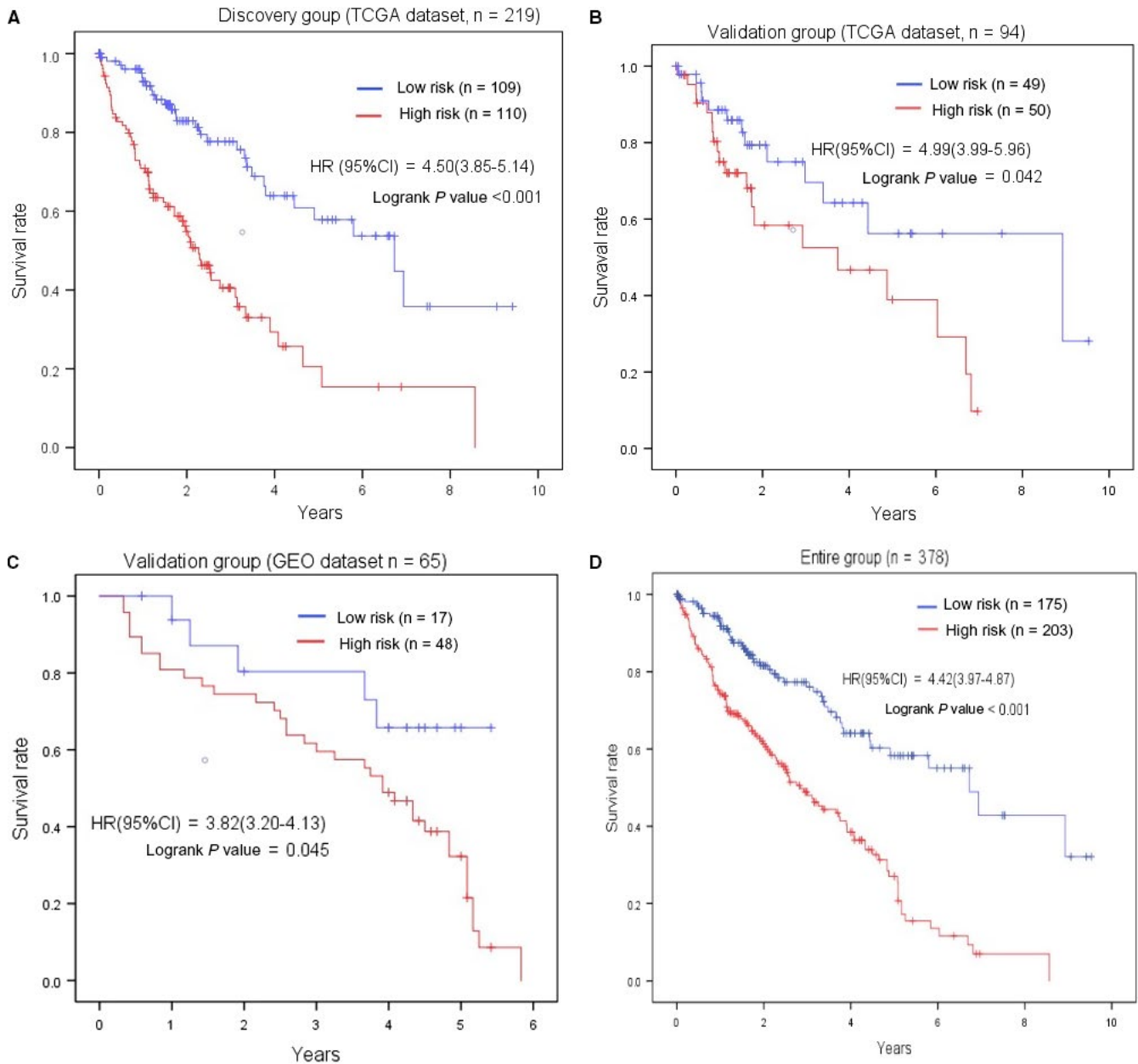


FIGURE 4 The association between five-lncRNA signature and overall survival in discovery and two validation groups. Kaplan-meier survival curves were plotted to estimate the overall survival probabilities for the low- and high-risk group in the discovery group (A). B, validation group from TCGA. C, validation group from GEO. D, entire group

and conducted a survival analysis (Figure 4D), which also confirmed the low-risk group has a better overall survival compared with high-risk group. In short, this finding further presents that the five lncRNAs are critical biomarkers which could affect the prognosis of patients with HCC.

3.5 | The five lncRNAs were associated with OS in patients with cancer at early stage

To explore the effects of five lncRNAs on clinical characteristics, we group patients based on five characteristics including

age, gender, grade, AJCC stage, and TNM staging. According to the results of analysis, we found the patients with cancer at early stage seem more related with the five lncRNAs (Table 2). As we observed from analysis, the stage I and N0 staging present statistical significance in discover group ($P < .05$). Moreover, we also explore the HR in different stage and found the trend is not statistical significance (P for trend $> .05$), which imply the risk model associated with five lncRNAs may only be related to OS in patients with early cancer. To validate our founding, we further calculated the prognostic values of the five lncRNAs in validation group and entire group (stage I, N0

TABLE 2 The association between five lncRNAs and overall survival of patients in discovery and validating groups

	Discovery group			Validation group			Entire group		
	Number (high/low)	HR (CI 95%)	P values	Number (high/low)	HR (CI 95%)	P values	Number (high/low)	HR (CI 95%)	P values
Age									
<=60	49/45	5.08 (4.11-6.06)	.003	19/24	4.12 (3.07-5.17)	.268	68/69	4.83 (4.05-5.61)	.006
>60	47/56	4.06 (3.21-4.91)	<.001	20/25	6.12 (4.73-7.50)	.707	67/61	4.46 (3.86-5.38)	<.001
Unknown	14/7	2.44 (1.83-3.05)	.99	1/5			15/12	3.05 (2.22-3.89)	.897
Gender									
Male	60/74	4.67 (3.90-5.44)	<.001	32/30	4.66 (3.80-5.52)	.898	92/104	4.69 (4.03-5.35)	.001
Female	50/34	4.02 (3.15-4.89)	.001	8/24	5.37 (3.74-7.00)	.766	58/58	4.47 (3.65-5.28)	.001
Grade									
G1	7/23	4.88 (3.71-6.05)	.001	6/8	4.17 (2.76-5.59)	.745	13/31	4.17 (3.72-5.70)	.014
G2	47/50	5.01 (4.02-6.01)	<.001	15/23	5.27 (3.86-6.69)	.460	70/66	5.10 (4.29-5.92)	.003
G3	36/24	3.98 (2.81-5.14)	.002	16/16	5.91 (4.14-7.69)	.187	52/40	4.95 (3.88-6.02)	.001
G4	20/11	2.78 (2.06-3.50)	.707	3/7			23/18	2.98 (2.32-3.65)	.714
AJCC stage									
I	45/58	5.24 (4.38-5.11)	.007	21/17	6.31 (4.46-7.95)	.042	66/75	2.34 (1.34-4.09)	.003
II	20/26	4.42 (2.98-5.87)	.001	9/18	2.80 (0.50-6.44)	.362	34/39	2.15 (1.09-4.25)	.027
III	30/17	3.59 (2.35-4.82)	.083	9/10	1.84 (0.28-2.50)	.764	22/44	2.09 (1.02-4.29)	.043
IV	15/7	2.51 (1.73-3.29)	.258	5/4	3.07 (0.09-12.60)	.955	12/19	2.02 (0.69-5.88)	.195
P for trend			>.05			>.05			>.05
TNM									
T									
T1	47/62	5.16 (4.33-6.00)	.005	21/19	3.25 (2.53-6.81)	.035	68/81	5.80 (4.90-6.71)	.006
T2	25/26	4.05 (2.78-5.32)	.001	11/20	5.86 (4.10-7.61)	.322	37/45	4.87 (3.80-5.93)	.039
T3	31/15	3.71 (2.47-4.94)	.101	9/11	3.26 (2.51-3.78)	.041	40/26	3.39 (2.53-4.26)	.151
T4	7/5	1.41 (0.86-1.96)	.301	5/6	2.98 (2.01-3.94)	.814	12/11	1.49 (1.01-1.96)	.518
P for trend			>.05			>.05			>.05
N									
N0	77/70	5.21 (4.37-6.04)	<.001	30/33	5.47 (4.24-6.70)	.039	107/103	5.43 (4.73-6.13)	<.001
N1	33/38	3.13 (2.51-3.75)	.042	14/18	4.53 (3.27-5.78)	.450	53/48	3.41 (2.82-3.99)	.241
M									
M0	77/81	4.99 (4.18-5.80)	<.001	26/39	5.58 (4.42-6.75)	.119	103/120	5.22 (4.54-5.89)	<.001
M1	33/27	3.22 (2.51-3.93)	<.001	14/15	3.33 (1.98-4.69)	.176	47/42	3.29 (2.66-3.92)	.028

staging). Kaplan-Meier analysis method was conducted to visualize the OS between high-risk and low-risk groups in patients with early cancer (Figure 5).

3.6 | Some pathways were associated with the five lncRNAs based on method of WGCNA

To further investigate the potential biological functions of the five lncRNAs, we used the WGCNA method to cluster genes based on the data of DEmRNA obtained from TCGA. We identified a total of nine coexpression modules with the threshold of five and found that blue module was positively correlated with the risk score (P value = .03), and GO functional biology processes and KEGG pathways enriched analysis using the genes enrolled into blue module are shown in Figure 6. GO functional biology processes include cell division, organelle fission, cytoskeletal part, and so on. KEGG pathway enrichment analysis was then performed on the basis of the package (clusterProfiler) of R. Four pathways enriched by DEmRNAs from ceRNA network include Cell cycle, p53 signaling pathway, Oocyte meiosis, and progesterone-mediated oocyte maturation, which were related to prognosis of HCC.

3.7 | The five lncRNAs were associated with immune cells infiltration

To further explore DElncRNA, we developed CIBERSORT method to search the most significant tumor-infiltrating immune cells and its correlation with immune cell type in liver cancer related to the DEmRNA (34 mRNA). Based on the CIBERSORT, we found the five lncRNAs played critical in the enumeration and activation status of five immune cell subtypes between paired cancer and normal tissue. Figure 7A summarizes the results obtained from 253 patients. There existed significant variation between normal and tumor group, which indicated the different subpopulations were closely correlation. Compared to normal tissue, HCC tissue related to five lncRNAs contained a higher proportion for T CD4⁺ memory activated, NK cells activated and mast cells activated ($P < .05$), while the monocytes and neutrophils decreased (Figure 7C). These results indicated that this five lncRNAs played critical role in immune infiltration and the characters with a tightly regulated process may have important clinical meanings. We explored the relationship of risk score with the significant immune cell based on five lncRNAs, the results also confirmed the five lncRNAs has close relation to the five types of immune infiltration cell (Figure 8). Besides, to estimate the accuracy, we also use the validation group to invalidate the results (Figure 8F-J).

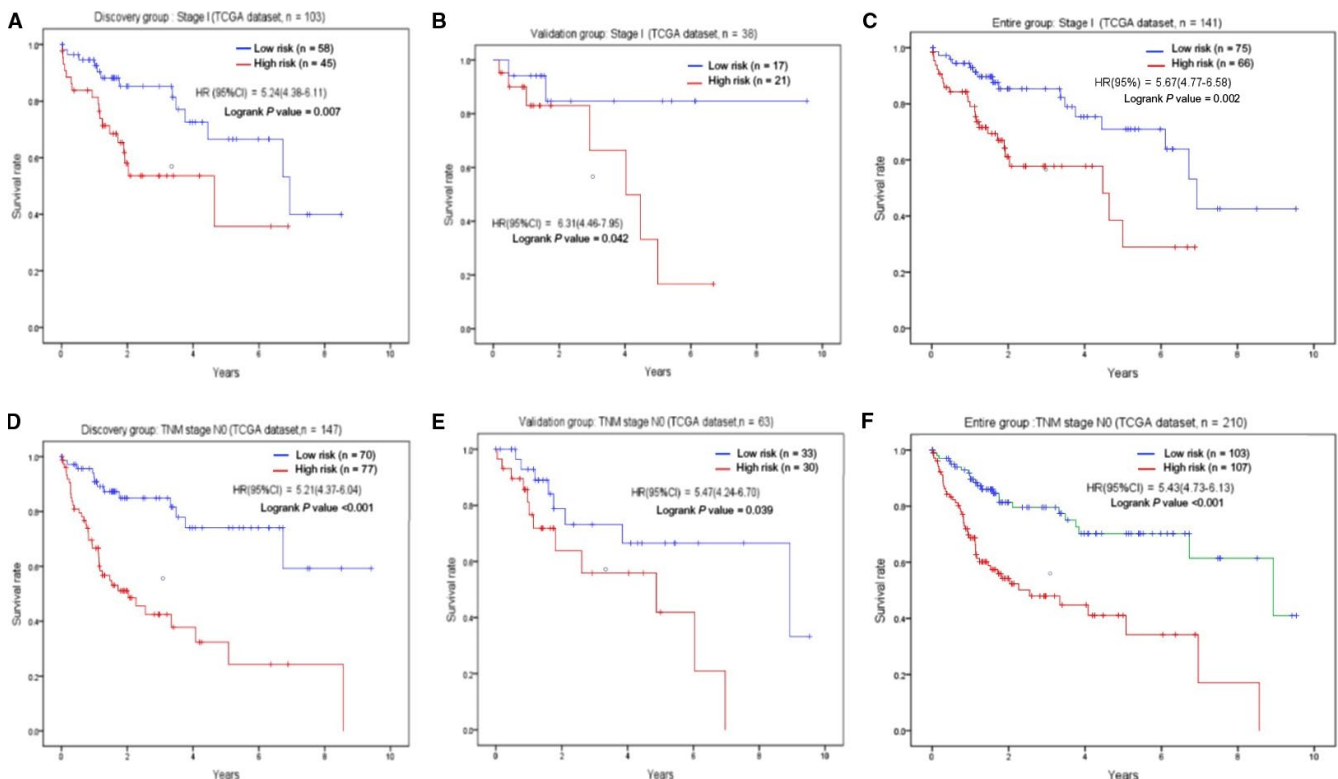


FIGURE 5 The association between five-lncRNA and overall survival in patients with early cancer. The survival curves of discovery group with stage I (A), validation group (B), entire group (C) were plotted. The similar results were obtained from the patients with N0 in discovery group (D), validation group (E), entire group (F)

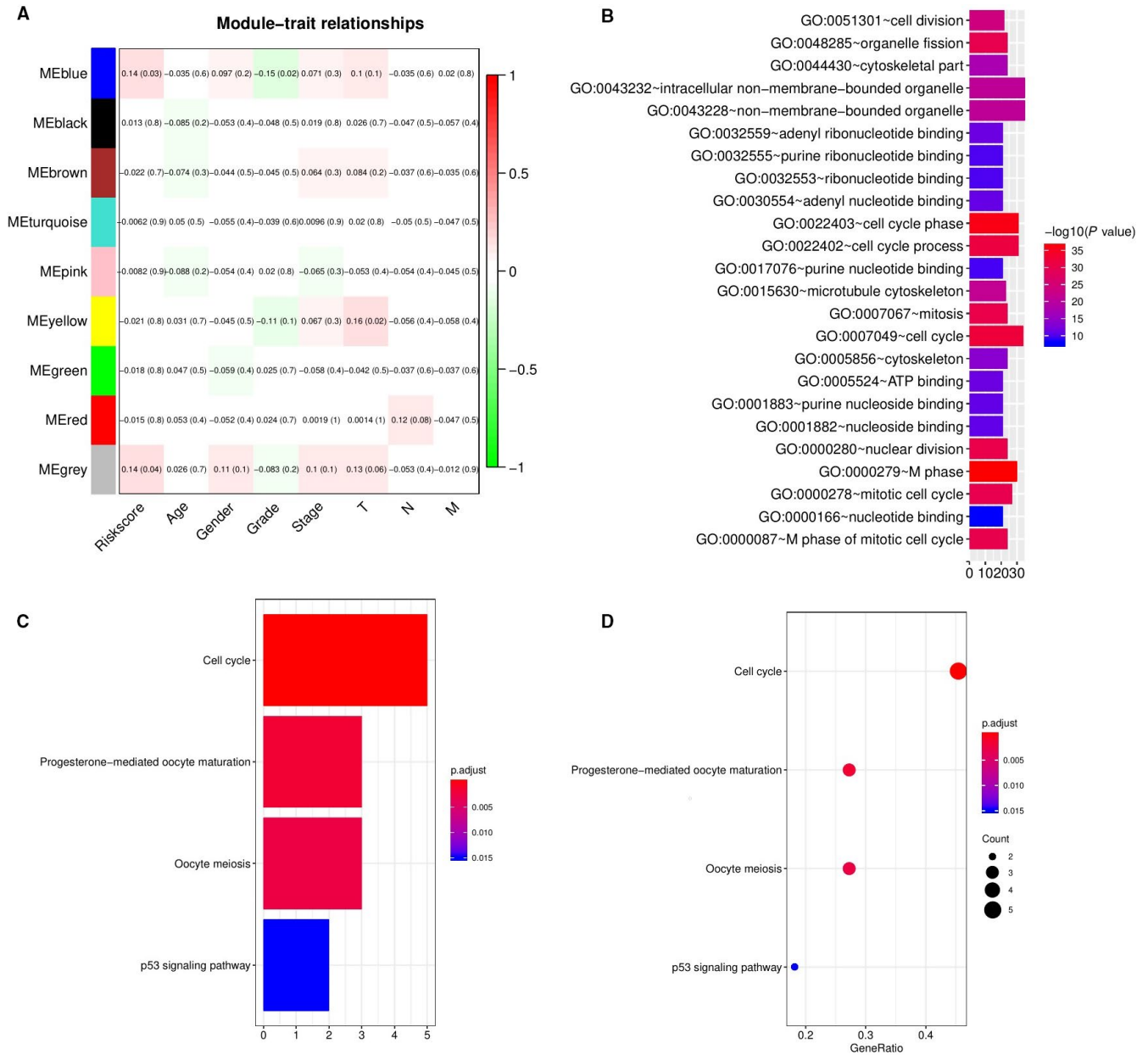


FIGURE 6 WGCNA predicted GO and KEGG pathways associated with the five-lncRNA signature. A, the gene clusters obtained by WGCNA method. B, the GO analysis of the co-expressed genes in blue module. C and D, significantly enriched pathways of the genes in blue module

3.8 | Transcription regulation prediction of ceRNA in HCC

We first got the promoter region of the five lncRNAs, and found most of lncRNAs was located in sense strand except one lncRNA LINC00162 which present reverse transcription. The promoter region was calculated based on transcription direction, respectively. Then, the TFBS were predicted in the promoter region using the web tool “UCSC” and “JASPAR.” Figure 9 present the TFBS of five lncRNAs. As known, the TFBS which present the same transcription direction with target gene has the higher

value to explore. We further explore the more significant TFBS, we got a total 96 kinds of TFBS, of which 20 kinds of TFBS were achieved in AATK-AS1 promoter region, 23 kinds of TFBS in C10orf91 promoter region, 28 kinds of TFBS in LINC00162 promoter region, 13 kinds of TFBS in LINC00200 promoter region, and 12 kinds of TFBS in LINC00501 promoter region (Figure 9). To better understand the mechanisms of TFBS, we conducted a KEGG pathways enriched analysis based on these TFBS associated with the five lncRNAs, and found these TFBS play critical role in transcriptional misregulation in cancer, Endocrine resistance, cGMP–PKG signaling pathway, and

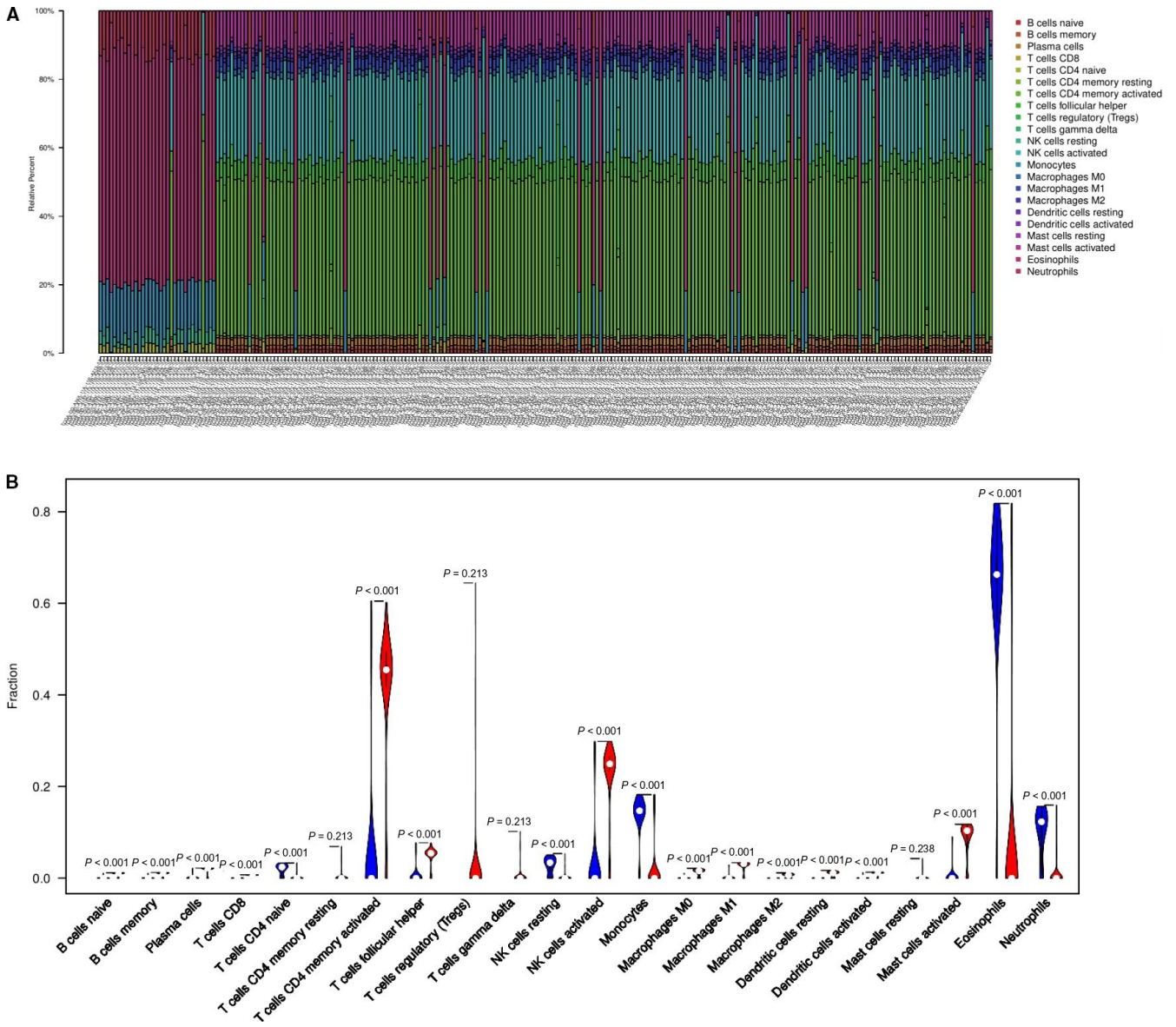


FIGURE 7 Immune cell composition in liver cancer and healthy livers. A, composition of infiltrating immune cells in different patients. B, violinplot visualizing the differentially infiltrated immune cells in cancer and normal groups on the basis of the 34DEmRNA

so on (Figure 10). These findings predicted the potential TFBS and attributed to understand the mechanisms of the ceRNA network with TFBS in HCC.

4 | DISCUSSION

The novel hypothesis has been confirmed that each RNA sharing the same MREs could interact with or compete each other, which present a new mechanism of gene expression regulation that could be used to further understand of various diseases including cancer. In the present study, we identified DEMiRNAs, DELncRNAs, and DEMRNA between HCC tissues and normal tissues, we then construct a

lncRNA-miRNA-mRNA ceRNA network the DELncRNAs in the ceRNA were analyzed for their association with the survival and clinical features of HCC patients, and we conducted validated the association with OS in two independent datasets. Besides, we found the expression of these five lncRNAs presented a statistical significance in patients with early stage.

In our study, we obtained a total five lncRNAs associated with clinical characters in HCC, including AATK-AS1, C10orf91, LINC00162, LINC00200, and LINC00501. Some of these lncRNAs were reported in cancer for first time including AATK-AS1 and LINC00200, while C10orf91 LINC00162, LINC00501, and HTR2A-AS1 has been reported in other type of cancer. Recent studies reported that

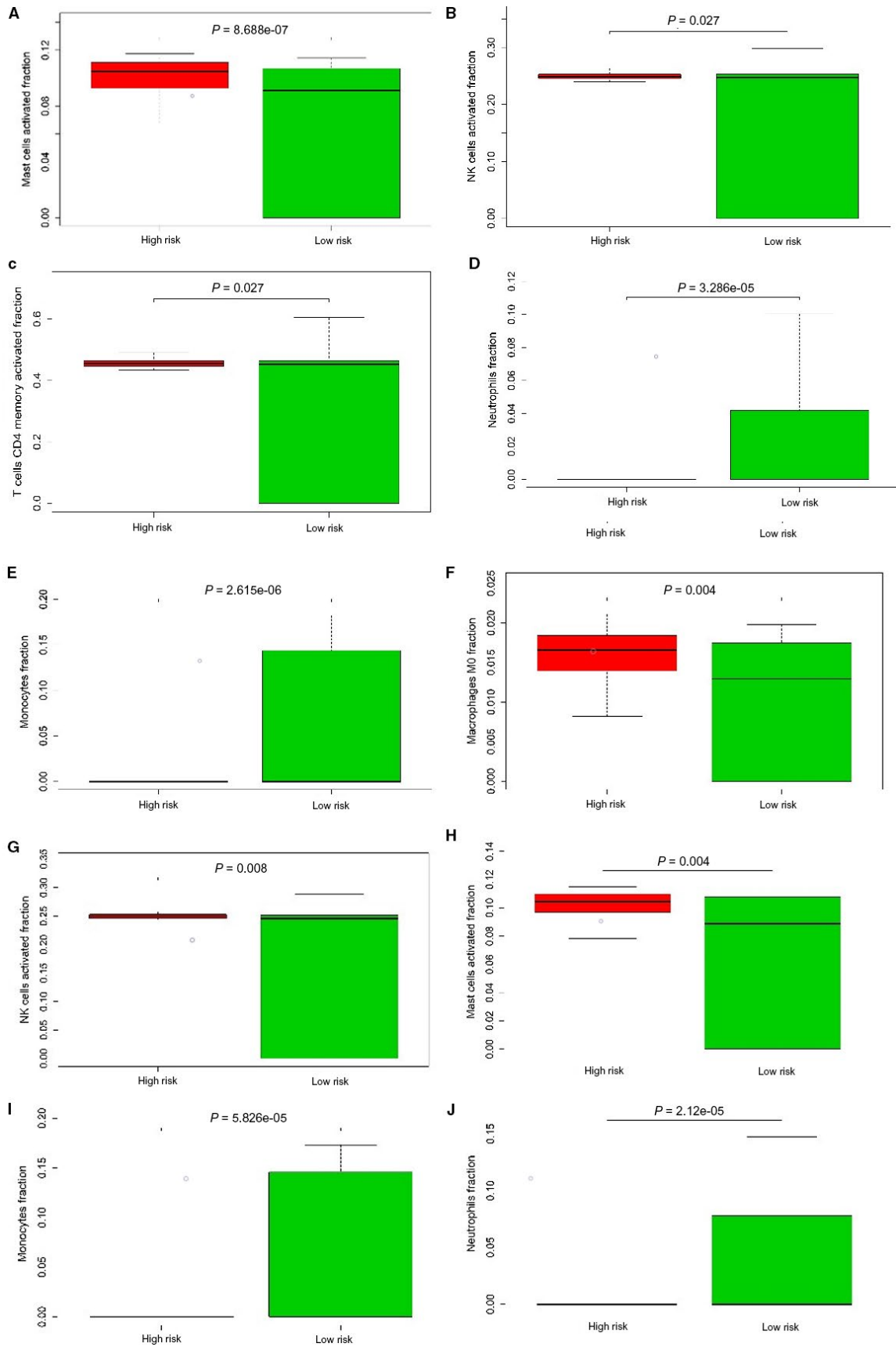


FIGURE 8 The relation of the immune cells to the five-lncRNAs. A-C, CD4⁺ memory activated, NK cells activated and mast cells activated accounted for a higher proportion in high risk group. D and E, Monocytes and neutrophils decreased. In validation group, we obtain the same results (F-J)

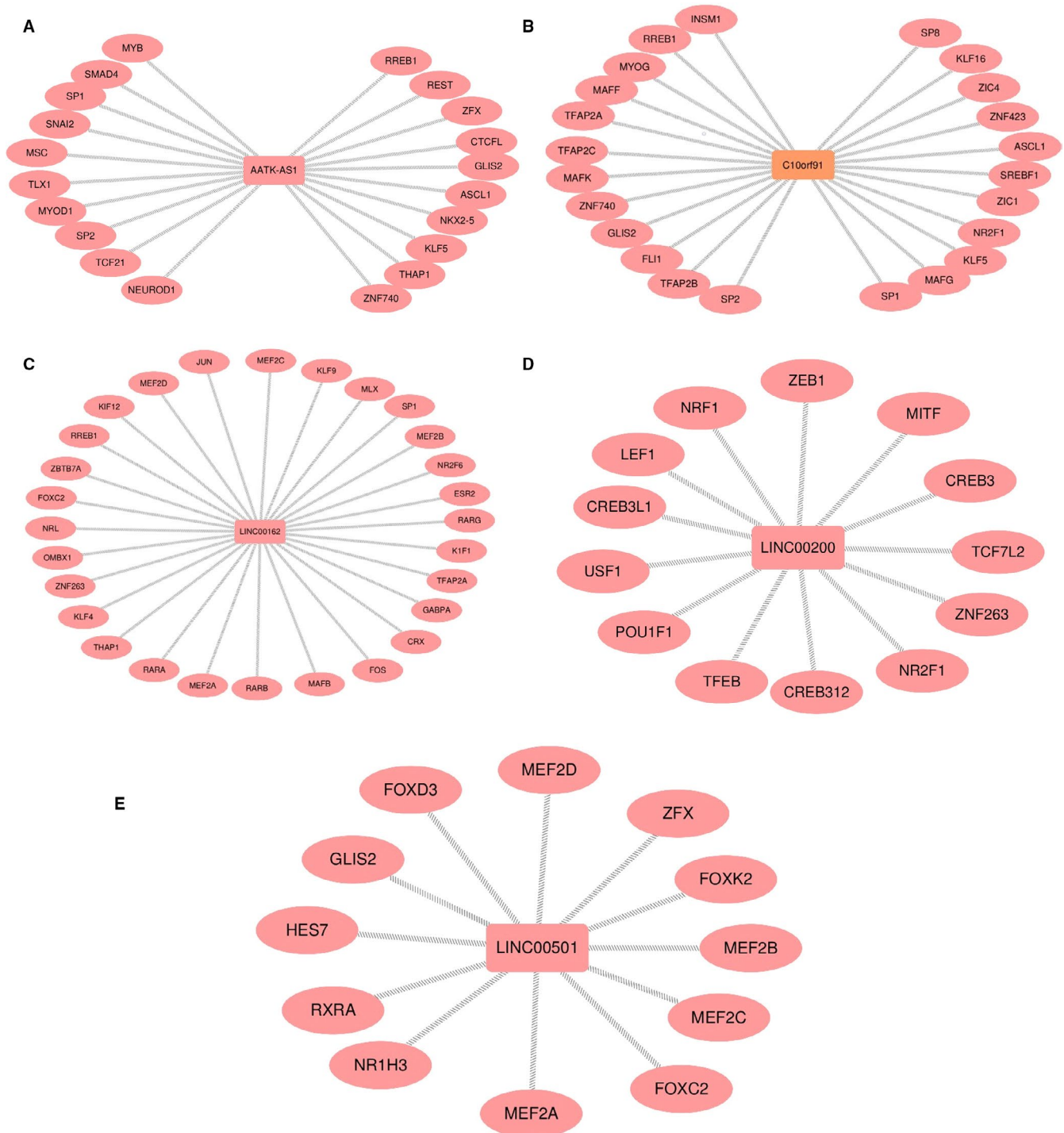
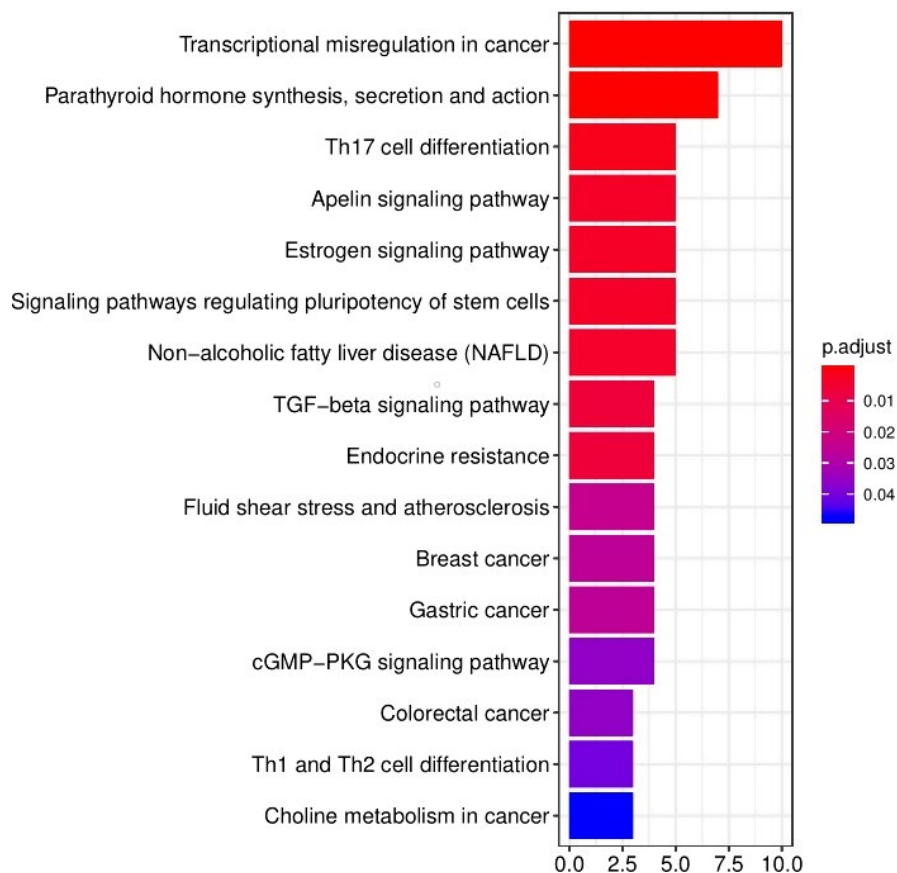


FIGURE 9 The potential TFBS of five lncRNA were predicted based on the same transcription direction

C10orf91 were related with OS on the basis of complex integrated analysis of lncRNAs-miRNAs-mRNAs regulatory network in oral squamous cell carcinoma,²⁰ LINC00162 named p38 inhibited cutaneous squamous cell carcinoma associated with lncRNA which promotes growth of cutaneous squamous cell carcinoma by regulating ERK1/2 activity,²¹ the rest of lncRNAs were all related to cell proliferation and contribute to carcinogenesis.²²⁻²⁴

One of the important findings in this study was that the five lncRNAs were associated with OS in early cancer stages. Our data showed their association reached a high significance in discovery group and in validation group, which imply that the five lncRNAs were a significant prognostic symbol for HCC with early stage. In order to investigate the potential mechanisms on the five lncRNAs, we further conducted bioinformatic analysis. We clustered

FIGURE 10 We conducted a KEGG analysis based on the TFBS associated with the five lncRNA



nine gene modules from 1987 differentially expressed genes based on the method of WGCNA, and found the blue modules was associated with the five lncRNAs. Pathway enrichment analysis further suggested that genes enrolled into blue module were mostly enriched cell cycle and p53 signaling pathway which contributes to enhanced proliferation of breast cancer cells,²⁵ indicating that the five lncRNAs could affect cell and consequently contributed to tumor progression. In other word, we may conclude that the five lncRNAs regulated the growth of HCC by regulating p53 signaling pathway. Of course, there still need more further studies to confirm. In short, the five lncRNAs play a major role in influencing prognosis.

The tumor-related microenvironment including immune cells, fibroblasts, and endothelial cells could make inhibitory effect on malignant cell, but with progression, tumor cells could grow, invasion, even metastasis by circumvent inhibitory signals and immune cells.²⁶ There is a complex relation of immune cells and malignant cells in cancer which has high relevance to immune system in either tumor-promoting or tumor-inhibiting roles, in present study, we infer the proportions of 22 immune cell from the DEMRNA used to construct ceRNA network using a silicon analysis, known as CIBERSORT.²⁷ Then, we conducted comprehensive analysis of clinical impact of immune response in HCC. We further compared the immune cells in high-risk and low-risk group. CIBERSORT was

always used based on thousands of genes screened from raw data, which means that the hub gene is not clear.^{19,28} Aimed at DEMRNA enrolled into ceRNA network, our data reveal that CD4 memory activated, NK cells activated and mast cells resting compared with normal increased and present statistically significance, while monocytes and neutrophils decreased. Thus, it can be seen that CD4 memory activated cells play a role in the development of HCC, which confirmed that the five lncRNAs have closely relation to immune cells and be a potential therapeutic target. Based on the above immune cells associated with the DEMRNAs, the high-risk present significant difference compared with low-risk group, which could validate the relation of five lncRNAs with immune cells.

We could better understand the potential mechanism of the five lncRNAs through the way of predicting the TFBS. The promoter region of the five lncRNAs was predicted using the web tool “NCBI,” then, we can obtain the potential TFBS of the five lncRNA based on the promoter region. The TFBS was screened to further explore which has the same transcription direction to the lncRNA. KEGG enrich pathways were conducted on the screened TFBS to explore further functions. The top pathway was transcriptional misregulation in cancer, and the result was also reported in Lee’s study that misregulation of gene expression can cause a broad range of diseases including cancer and participated in cancer progression by regulating cell cycle

and cell proliferation.²⁹ TGF-beta signaling pathway was an important role in cervical cancer.³⁰ Another study reported suppressing the TGF-beta signaling pathway was a well-known immunosuppressor and proangiogenic factor, contributing to fight against cancer.³¹ These studies have reported these function or pathway which might explain the potential mechanism of HCC and provide the further thought to conduct. We assumed that the TFBS might play important roles in regulating the ceRNA in liver cancer.

To conclude, lncRNA play critical roles in the development of cancer and may have close relation to prognosis. We constructed a ceRNA network, and a risk-score model based on five lncRNAs to predict the OS of HCC patients, which could help people to assess the prognosis of HCC with higher accuracy. The risk association of five lncRNAs was more significant in patients with early stage. We explored the immune cell associated with the five lncRNAs, which contributed to immune therapy. Prediction of TFBS associated with five lncRNAs help us to understand the potential mechanism. It still needs further studies to validate our founding.

ACKNOWLEDGMENTS

The Natural Science Foundation of Tianjin (16JCYBJC27100).

CONFLICT OF INTERESTS

The authors declare no competing interests.

AUTHOR CONTRIBUTIONS

Xin Lou was involved in writing of this article, data processing, and preparing of figures. Jun Li was involved in data processing. Dong Yu was involved in preparing of figures. Ya-Qing Wei and Shuang Feng were involved in writing of this article. Jin-Jin Sun was involved in putting forward the idea. All authors have read and approved the final manuscript.

ORCID

Jin-Jin Sun  <https://orcid.org/0000-0003-3676-1437>

REFERENCES

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *Cancer J Clin.* 2019;69:7-34.
- Bruix J, Reig M, Sherman M. Evidence-based diagnosis, staging, and treatment of patients with hepatocellular carcinoma. *Gastroenterology.* 2016;150:835-853.
- El-Serag HB. Hepatocellular carcinoma. *N Engl J Med.* 2011;365:1118-1127.
- Llovet JM, Ricci S, Mazzaferro V, et al. Sorafenib in advanced hepatocellular carcinoma. *N Engl J Med.* 2008;359:378-390.
- Abou-Alfa GK, Meyer T, Cheng A-L, et al. Cabozantinib in patients with advanced and progressing hepatocellular carcinoma. *N Engl J Med.* 2018;379:54-63.
- Ramanathan M, Porter DF, Khavari PA. Methods to study rna-protein interactions. *Nat Methods.* 2019;16:225-234.
- Hendrickson DG, Hogan DJ, McCullough HL, et al. Concordant regulation of translation and mrna abundance for hundreds of targets of a human microRNA. *PLoS Biol.* 2009;7:e1000238.
- Bossi L, Figueroa-Bossi N. Competing endogenous RNAs: a target-centric view of small RNA regulation in bacteria. *Nat Rev Microbiol.* 2016;14:775-784.
- Seitz H. Redefining microRNA targets. *Curr Biol.* 2009;19:870-873.
- Paraskevopoulou MD, Georgakilas G, Kostoulas N, et al. Diana-lncbase: experimentally verified and computationally predicted microRNA targets on long non-coding rnas. *Nucleic Acids Res.* 2013;41:D239-D245.
- Wang KC, Yang YW, Liu BO, et al. A long noncoding RNA maintains active chromatin to coordinate homeotic gene expression. *Nature.* 2011;472:120-124.
- Chen Z, Liu H, Yang H, Gao Y, Zhang G, Hu J. The long noncoding RNA, TINCR, functions as a competing endogenous RNA to regulate pdk1 expression by sponging mir-375 in gastric cancer. *Onco Targets Ther.* 2017;10:3353-3362.
- Cui M, Wang J, Li Q, Zhang J, Jia J, Zhan X. Long non-coding RNA HOXA11-AS functions as a competing endogenous RNA to regulate ROCK1 expression by sponging miR-124-3p in osteosarcoma. *Biomed Pharmacother.* 2017;92:437-444.
- Liu Y, Guo G, Zhong Z, et al. Long non-coding RNA FLVCR1-AS1 sponges miR-155 to promote the tumorigenesis of gastric cancer by targeting c-Myc. *Am J Transl Res.* 2019;11:793-805.
- Deng L, Liu G, Zheng C, Zhang L, Kang Y, Yang F. Circ-LAMP1 promotes T-cell lymphoblastic lymphoma progression via acting as a ceRNA for miR-615-5p to regulate DDR2 expression. *Gene.* 2019;701:146-151.
- Fan Q, Liu B. Comprehensive analysis of a long noncoding RNA-associated competing endogenous RNA network in colorectal cancer. *Onco Targets Ther.* 2018;11:2453-2466.
- Xia T, Liao QI, Jiang X, et al. Long noncoding RNA associated-competing endogenous RNAs in gastric cancer. *Sci Rep.* 2014;4:6088.
- Lin T, Fu Y, Zhang X, et al. A seven-long noncoding RNA signature predicts overall survival for patients with early stage non-small cell lung cancer. *Aging.* 2018;10:2356-2366.
- Ali HR, Chlon L, Pharoah PD, Markowitz F, Caldas C. Patterns of immune infiltration in breast cancer and their clinical implications: a gene-expression-based retrospective study. *PLoS Medicine.* 2016;13:e1002194.
- Li S, Chen X, Liu X, et al. Complex integrated analysis of lncRNAs-miRNAs-mRNAs in oral squamous cell carcinoma. *Oral Oncol.* 2017;73:1-9.
- Piipponen M, Nissinen L, Farshchian M, et al. Long noncoding RNA PICSAR promotes growth of cutaneous squamous cell carcinoma by regulating ERK1/2 activity. *J Invest Dermatol.* 2016;136:1701-1710.
- Li M, Zhang W, Zhang S, Wang C, Lin Y. PART1 expression is associated with poor prognosis and tumor recurrence in stage I-III non-small cell lung cancer. *J Cancer.* 2017;8:1795-1800.

23. Hou Y, Yu Z, Tam NL, et al. Exosome-related lncRNAs as predictors of HCC patient survival: a prognostic model. *Am J Transl Res.* 2018;10:1648-1662.
24. Zhang H, Wang Z, Wu J, Ma R, Feng J. Long noncoding RNAs predict the survival of patients with colorectal cancer as revealed by constructing an endogenous RNA network using bioinformation analysis. *Cancer Med.* 2019;8:863-873.
25. Wan W, Hou Y, Wang K, Cheng Y, Pu X, Ye X. The LXR-623-induced long non-coding RNA LINC01125 suppresses the proliferation of breast cancer cells via PTEN/AKT/p53 signaling pathway. *Cell Death Dis.* 2019;10:248.
26. Quail DF, Joyce JA. Microenvironmental regulation of tumor progression and metastasis. *Nat Med.* 2013;19:1423-1437.
27. Newman AM, Liu CL, Green MR, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12:453-457.
28. Liu X, Wu S, Yang Y, Zhao M, Zhu G, Hou Z. The prognostic landscape of tumor-infiltrating immune cell and immunomodulators in lung cancer. *Biomed Pharmacother.* 2017;95:55?-61.
29. Lee TI, Young RA. Transcriptional regulation and its misregulation in disease. *Cell.* 2013;152:1237-1251.
30. Chen J, Deng Y, Ao L, et al. The high-risk HPV oncogene E7 up-regulates miR-182 expression through the TGF- β /Smad pathway in cervical cancer. *Cancer Lett.* 2019;460:75-85.
31. Fabregat I, Fernando J, Mainez J, Sancho P. TGF-beta signaling in cancer treatment. *Curr Pharm Des.* 2014;20:2934-2947.

How to cite this article: Lou X, Li J, Yu D, Wei Y-Q, Feng S, Sun J-J. Comprehensive analysis of five long noncoding RNAs expression as competing endogenous RNAs in regulating hepatoma carcinoma. *Cancer Med.* 2019;8:5735-5749. <https://doi.org/10.1002/cam4.2468>