Identification of Ten-lncRNA Signatures Predict Disease-free and Overall Survival for Hepatocellular Carcinoma

Si-shu Yang
West China Hospital of Sichuan University
Jiong Lu
West China Hospital of Sichuan University
Xian-ze Xiong (xiongxianze@scu.edu.cn)
West China Hospital of Sichuan University

Research Article

Keywords: long non-coding RNA, prognosis, risk score, hepatocellular carcinoma

Posted Date: December 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-613111/v2

Abstract

Background: The prognosis of hepatocellular carcinoma (HCC) is bleak though it has been improved over recent years. Early diagnosis could improve the survival. Plenty of researches indicate that long non-coding RNAs (lncRNAs) could play an important role in prognostic prediction of cancer as a kind of biomarker.

Results: We identified and validated ten-lncRNAs based signatures to predict disease-free survival (DFS) and overall survival (OS) of HCC respectively from lncRNA expression data of HCC patients in The Cancer Genome Atlas (TCGA) database. Stratified survival analysis showed that the performance of lncRNAs related signatures was better than tumor, node, metastasis(TNM) staging system. Functional enrichment analysis showed that organelle fission and regulation of mRNA metabolic process were significantly enriched in differentially expressed lncRNAs (DElncRNAs). Transcriptional misregulation in cancer and mitogen-activated protein kinase (MAPK) signaling pathway were significantly enriched pathways in the pathway enrichment analysis.

Conclusion: we constructed two lncRNAs based signatures which could predict prognosis of HCC more accurate than the traditional ways.

Background

HCC is a aggressive malignancy with high incidence and mortality[1]. Due to the high rates of metastasis and recurrence, the overall survival rate of HCC remains low though great efforts have been made in the diagnosis and treatments[2, 3]. Consequently, it is urgent to identify reliable molecular biomarkers that can predict early recurrence and prognosis of HCC to help physicians with decision-making.

Long non-coding RNAs are transcripts longer than 200 nucleotides that have no protein-coding potential, however, they play a vital role in lots of physiologic processes[4-6]. Over the last few decades, genome-wide sequencing facilitated the discovery of a series of lncRNAs. Recently, increasing evidences have indicated abnormal expression of lncRNAs is frequent in various cancers, and is closely linked to cancer initiation, development, progression, and metastasis[7-10]. Emerging studies suggested that lncRNAs might act as potential predictors for clinical outcomes in cancers, including HCC[11-14].

Currently, with the advancements of the big data era, high-throughput data and clinical data could be achieved by mining previously published public databases such as The Cancer Genome Atlas and The Gene Expression Omnibus (GEO). Researches have developed a couple of systems biology methods to model lncRNA-based signatures in order to predict prognosis more accurately[15-18]. But in fact, the lncRNA signatures that can effectively predict the recurrence and prognosis of HCC as well as their prognostic value have not yet been fully explored.

In this study, we screened prognosis-associated lncRNAs and established a prognostic risk assessment model through analyzing lncRNA expression records of 371 HCC patients from TCGA database. Two ten-lncRNA signatures for disease-free survival and overall survival respectively were constructed and successfully validated. It turns out these lncRNA—based models could be served as dependable prognostic predictors for HCC.

Results

Identification of prognosis-associated DElncRNAs

The workflow of the study is displayed in Figure 1. As shown in Fig 2A, B, a total of 232 DElncRNAs with |logFC| >1.5 and padj < 0.05 were identified by R package 'limma'. Of them, 191 lncRNAs were found to be upregulated and 41 to be downregulated, which indicated that these DElncRNAs could distinguish HCC from normal tissues. In addition, to determine significant prognosis-associated lncRNAs (p<0.01) in HCC, univariate Cox regression was performed. In total, 2113 lncRNAs were chosen for further analyses. Among them, 1193 and 920 lncRNAs were significantly correlated with DFS and OS, respectively. Following this, 60 DFS-related and 45 OS-related DElncRNAs were selected for the next step of the analysis (Fig 2C, D).

Construction of lncRNAs signature

LASSO regression was adopted to build the risk score model for prognosis prediction in the training cohort at a 20-fold cross-validation (Fig 3). 10 lncRNAs related to DFS and 10 lncRNAs related to OS were selected by lambda.min value. The expression levels of these prognostic lncRNAs and their coefficients were constructed signatures respectively. The disease free survival risk score = (0.13915 * AC009005.2) + (0.14837 * CASC9) + (0.26053 * CTC_338M12.5) + (0.13069 * DYNLL1_AS1) + (-0.06954 * LINC01018) + (0.08498 * PRRT3_AS1) + (0.13913 * RP11_488L18.10) + (0.22265 * RP11_739N20.2) + (0.19098 * RP11_977G19.5) + (0.06799 * WAC_AS1). The overall survival risk score = (0.10825 * AC007405.6) + (-0.12949 * CTC_297N7.9) + (0.17571 * CTD_2510F5.4) + (-0.07341 * F11_AS1) + (-0.00409 * LINC00152) + (0.31658 * LINC01138) + (0.20993 * PCAT6) + (0.21144 * PRRT3_AS1) + (0.25757 * RP11_307C12.11) + (0.31922 * RP11_479G22.8).

Validation and confirmation of prognostic performance of the lncRNAs signature

Using those formulas to figure up the risk score of each patient. All patients were classified into high-risk and low-risk groups in the light of the cut off value. Kaplan Meier curves showed patients with high-risk scores had shorter DFS (P<0.0001) and OS (P<0.0001) than those with low-risk scores in the training group. Similar results were observed in the validation and entire set (Figure 4).

In addition, we analyzed the clinicopathological characteristics between the high-risk and low-risk group in training, validation and entire set respectively by chi-square test. Detailed information was showed in Table 1. For the DFS-related group, the distribution of grade between high-risk and low-risk group showed a significant difference in the training set(P=0.002). In the validation set, there were statistically significant differences in pT(P=0.008), stage(P=0.006) and grade(P=0.004) between high-risk and low-risk group. For the OS-related group, similar results were obtained in the entire set. There were significant differences in pT(P=0.002), stage(P=0.004) and grade(P<0.001) between high-risk and low-risk group. The proportion of patients with pT3-pT4, stage III-IV and grade 3-4 were higher in the high-risk group than low-risk group. On the other hand, for the DFS-related cohort and OS-related cohort, patients with pT3-pT4, stage III-IV and grade 3-4 had higher risk scores than the patients with pT1-pT2, stage I-II and grade 1-2 in the entire set (Figure 5).

Univariate and multivariate Cox regression were adopted with the lncRNA-based signature and clinicopathological characteristics. The outcome was presented in Figure 6. The lncRNA-based signature was an independent

predictive factor of DFS and OS for training and entire sets (P<0.05). In the univariate analyses, pT and stage of the tumor were demonstrated to be risk factors for both DFS and OS of hepatocellular carcinoma patients in training, validation and entire cohorts (P<0.05). Pathological metastasis was also a significant risk factor for DFS and OS in the univariate regression (P<0.05).

The performance of the 10-lncRNA signature was evaluated by the area under the ROC curve. The area under curve (AUC) value of the 1,3,5-year DFS predicted in the entire set was 0.73, 0.69 and 0.68 respectively (Figure 7A). The AUC value of the 1,3,5-year OS predicted in the entire set was 0.76, 0.72 and 0.71 (Figure 7B). All of the AUC values showed the good performance of lncRNA-based signature. To explore the prognostic value of lnc-RNA based signature, we compared the AUC value of lncRNA-based signature and American joint committee on cancer (AJCC) 8th TNM stage (Figure 8). The AUC value of TNM stage for DFS was 0.67 lower than the AUC value of lncRNA-based signature. However, the AUC value of TNM stage combined lncRNA-based signature was 0.79 better than the AUC value of TNM stage or lncRNA-based signature alone. Similar results were obtained for OS analysis, which suggested that the combination of TNM stage and lncRNA-based signature could improve the prognosis predictive ability.

What' more, in order to further analyze the applicable clinical characteristics of lncRNA-based signature, we divided all the patients into different subgroups by TNM stage integrating with lncRNA-based signature. Kaplan Meier curves were performed in the subgroups (Figure 9). The survival curves indicated that patients of stage I/II in the low-risk group had better survival rate than which in the high-risk group (P<0.0001), patients of stage III/IV in the high-risk group had worse survival prognosis than which in the low-risk group (P<0.0001).

Functional enrichment analyses

Firstly, 3212 and 250 target genes that were co-expressed with DFS/OS related lncRNAs were extracted from RAID 2.0 database. These significantly correlated genes were used for GO analysis and KEGG enrichment analysis to determine the potential mechanism of the DFS/OS related lncRNAs in regulating HCC (Figure 10). GO enrichment analysis found the DFS related DElncRNAs were mainly enriched in organelle fission, nuclear division mitotic nuclear division, and chromosome segregation, etc. Moreover, regulation of mRNA metabolic process, regulation of RNA stability and pri-miRNA transcription by RNApolymeraseII were the most enriched function with OS related DElncRNAs. KEGG analysis showed that the target gene of DFS related DElncRNAs had significant link to lots of enriched pathways, including cell cycle, p53 signaling pathway, tumor necrosis factor signaling pathway, and MAPK signaling pathway. What's more, the KEGG analysis of OS related DElncRNAs revealed that their targeted genes were involved in transcriptional misregulation in cancer, human T-cell leukemia virus 1 infection, interleukin-17 signaling pathway, cell cycle, and transforming growth factor–beta signaling pathway.

Discussion

HCC has high clinical and genic heterogeneity with high recurrent rate. Effective prognostic biomarkers are useful for physicians to take appropriate treatments to extend the disease free survival and overall survival. Recently several researches discovered non-coding RNA acted on carcinogenesis and tumor progression, including proliferation, metastasis, metabolic regulation and drug resistance, etc[19-24]. Lots of dysregulated lncRNAs show great potential as biomarker in prognostic evaluation of cancer. Several lncRNA-based prognostic models have been used for many malignancies, such as carcinoma of lung, bladder and breast[25-29], etc. Several researches

have constructed lncRNA-based risk models of HCC, but the existing risk models mostly based on the overall survival rather than disease free survival[30-33]. In this study, we screened both DFS-associated and OS-associated DElncRNAs in the TCGA databases by univariate Cox and LASSO regression, then established two lncRNA-based signatures focusing on DFS and OS respectively, both of the prognostic models were validated and showed good performance.

Several studies suggested that the expression level of LINC 00152 was increased in HCC and reported to be a great diagnostic lncRNA with good performance[34, 35]. LINC 00152 was found enhancing migration and invasion of tumor cells by neuro trophin receptor kinase (NTRK)3 and promoting proliferation of tumor by activating the mammalian target of rapamycin (mTOR) pathway [36, 37]. Except LINC 00152, rest of lncRNAs were the first time to be reported as prognostic factors of HCC as far as we knew. GO and KEGG enrichment analyses were conducted to figure out molecular mechanism of DElncRNAs related with DFS/OS. GO enrichment analysis showed that organelle fission and regulation of mRNA metabolic process were the most enriched functions of DFS and OS respectively. KEGG analysis suggested that the most enriched pathways of DFS and OS were MAPK signaling pathway and transcriptional misregulation in cancer respectively. These results indicated that these lncRNAs may involve in carcinogenesis and tumor development by those biological functions and pathways. However, the detailed mechanism still needs further experiments to demonstrate.

TNM stage is the most commonly used prognostic evaluation method during the clinical practice. We contrasted the performance of lncRNA-related model and TNM stage by AUC value and Kaplan Meier curve, demonstrating that lncRNA-based signatures had better performance in the prognostic prediction. Precise stratified analysis can help physicians make appropriate clinical decisions and is beneficial to improve prognosis at some extent. Recently, several studies suggested that patients received liver transplantation had better overall survival and lower recurrence rate than which treated with hepatectomy in early HCC[38-40]. Thus, it's crucial to select applicable candidates because of the shortage of donors. In the Kaplan Meier analysis, the survival of patients in stage I/II with low-risk were better than which with the high-risk ($P<0.0001$), which may be helpful with identifying the appropriate candidates of liver transplantation. However, patients of stage III/IV with high-risk had worse survival than which with low-risk($P<0.0001$), which may need more aggressive treatment strategies and more closely follow-up.

Our study has several limitations. Primarily, entire molecular mechanism of these lncRNAs in HCC development remains unknown, which needs further experiments to figure out. Secondly, part of momentous clinicopathological features associated with prognosis were not included in this study because of absence in the TCGA dataset, combining with clinicopathological features may improve the accuracy of the signatures. Finally, this is a retrospective research from TCGA database, the performance of these lncRNA-based signatures need to be validated by a larger scale external cohort.

Conclusions

We built two 10-lncRNA related prognostic model for DFS and OS respectively by univariate Cox and LASSO regression which stratified HCC patients more precise than TNM staging system and could be useful tools to provide accurate prognostic information for clinical physician with decision-making. However, there remains the needs of larger-scale studies to validate the prognostic accuracy of the signatures and the potential molecular functions of these lncRNAs.

Methods

Data acquisition

The lncRNA expression data and clinicopathological characteristics of HCC patients were downloaded from the TCGA database (https://xenabrowser.net/datapages), including 371 HCC tissues and 50 adjacent non-tumor tissues. 371 patients with available survival information were enrolled in the study totally. Clinicopathological parameters of the HCC patients are shown in Table 2. The ratio of training sets to validation sets was 2:1.

Data processing

DElncRNAs were screened by edgeR and limma packages from the entire lncRNA expression profiling data between tumor and normal samples. The threshold was set at |log2 fold change (FC)|>1.5 and p adj < 0.05. Meanwhile, univariate Cox regression was conducted to screen out lncRNAs that were significantly associated with DFS and OS using the survival package in R. The significance of the identified lncRNAs was tested by log-rank test, and p<0.01 was considered statistically significant.

Construction of lncRNAs signature

The significant lncRNAs were selected into the LASSO[41] regression analysis at 20-fold cross-validation to get the correlation coefficients. By linearly combining the expression level of selected lncRNAs weighted by the estimated coefficients, the following risk-score formula was constructed for calculation: Risk Score = exp_{gene1} * β_{gene1} + exp_{gene2} * β_{gene2} +⋯+ exp_{genen}* β_{genen} (exp, expression level, β, the regression coefficient in the LASSO regression analysis). According to the cut-off of the median risk score, HCC patients were classified into high- and low-risk group.

Confirmation of lncRNAs signature

Comparison of the differences of DFS and OS between high-risk and low-risk group was performed by Kaplan-Meier survival curves and log-rank tests in all sets. Differentiation of clinicopathological characteristics between high-risk group and low-risk group was evaluated by Pearson's χ^2 test. Univariate and multivariate Cox regression analyses were adopted to screen out the independent prognostic factors in the lncRNA-related signature and other clinical parameters. ROC were plotted to assess the specificity and sensitivity of lncRNA classifier in HCC prognostic prediction. All the analyses were performed by R software with the following packages: 'edgeR', 'glmnet', 'survival' and 'survminer'. A P value less than 0.05 was considered statistically significant.

Functional enrichment analyses

To screen target genes of DElncRNAs, human lncRNA-mRNA interaction was download from redundant arrays of independent disk (RAID) 2.0 database (www.rna-society.org/raid/). The correlated mRNAs were evaluated by functional enrichment analysis to explore the functions of the prognostic DElncRNAs using the clusterProfiler.

Significantly enriched gene ontology (GO) terms and kyoto encyclopedia of genes and genomes (KEGG) pathways with P value <0.05 were visualized using the R software.

Abbreviations

HCC – Hepatocellular carcinoma

LncRNAs – long non-coding RNAs

TCGA database – The Cancer Genome Atlas database

LASSO regression – least absolute shrinkage and selection operator regression

ROC curve – receiver operating characteristics curve

mRNA – messenger RNA

DFS – disease-free survival

OS – overall survival

TNM staging system – tumor, node, metastasis staging system

DElncRNAs – differentially expressed lncRNAs

MAPK signaling pathway – mitogen-activated protein kinase

GEO database – The Gene Expression Omnibus database

RAID – redundant arrays of independent disk

GO – gene ontology

KEGG – Kyoto encyclopedia of genes and genomes

AUC – area under curve

AJCC – American joint committee on cancer

NTRK – neuro trophin receptor kinase

mTOR – mammalian target of rapamycin

Declarations

Ethics approval and consent to participate

The study was carried out in accordance with relevant guidelines and regulations and approved by the ethics committee of west china hospital, Sichuan university.

Consent for publication

Not applicable.

Availability of data and materials

The datasets analyzed during the study are available in the http://xenabrowser.net /datapages and www.rna-society.org/raid/.

Competing interests

There are no conflicts to declare.

Funding

This study is supported by National Natural Science Foundation of China (No. 82002578) and Major Science and Technology Project of Sichuan Province, China (No. 2019YFS0041).

Authors' contributions

SS Yang and J Lu analyzed the data and wrote the manuscript; XZ Xiong designed and supervised the writing of the manuscript; all authors approved the final manuscript as submitted.

Acknowledgements

None.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A: Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin 2018, 68(6).
2. Forner A, Llovet JM, Bruix J: Hepatocellular carcinoma. Lancet (London, England) 2012, 379(9822).
3. Maluccio M, Covey A: Recent progress in understanding, diagnosing, and treating hepatocellular carcinoma. CA: a cancer journal for clinicians 2012, 62(6).
4. Bhan A, Soleimani M, Mandal SS: Long Noncoding RNA and Cancer: A New Paradigm. Cancer research 2017, 77(15).
5. Kopp F, Mendell JT: Functional Classification and Experimental Dissection of Long Noncoding RNAs. Cell 2018, 172(3).
6. Uszczynska-Ratajczak B, Lagarde J, Frankish A, Guigo R, Johnson R: Towards a complete map of the human long non-coding RNA transcriptome. Nature reviews Genetics 2018, 19(9).

7. Fatica A, Bozzoni I: Long non-coding RNAs: new players in cell differentiation and development. Nature reviews Genetics 2014, 15(1).
8. Yang F, Huo X-s, Yuan S-x, Zhang L, Zhou W-p, Wang F, Sun S-h: Repression of the long noncoding RNA-LET by histone deacetylase 3 contributes to hypoxia-mediated metastasis. Mol Cell 2013, 49(6).
9. Fernández-Barrena MG, Perugorria MJ, Banales JM: Novel lncRNA T-UCR as a potential downstream driver of the Wnt/β-catenin pathway in hepatobiliary carcinogenesis. Gut 2017, 66(7).
10. Li W, Zhang Z, Liu X, Cheng X, Zhang Y, Han X, Zhang Y, Liu S, Yang J, Xu B et al: The FOXN3-NEAT1-SIN3A repressor complex promotes progression of hormonally responsive breast cancer. J Clin Invest 2017, 127(9).
11. Yin D, Lu X, Su J, He X, De W, Yang J, Li W, Han L, Zhang E: Long noncoding RNA AFAP1-AS1 predicts a poor prognosis and regulates non-small cell lung cancer cell proliferation by epigenetically repressing p21 expression. Molecular cancer 2018, 17(1).
12. Wang Y, Li W, Chen X, Li Y, Wen P, Xu F: MIR210HG predicts poor prognosis and functions as an oncogenic lncRNA in hepatocellular carcinoma. Biomed Pharmacother 2019, 111
13. Lu M, Liu Z, Li B, Wang G, Li D, Zhu Y: The high expression of long non-coding RNA PANDAR indicates a poor prognosis for colorectal cancer and promotes metastasis by EMT pathway. J Cancer Res Clin Oncol 2017, 143(1).
14. Lee YR, Kim G, Tak WY, Jang SY, Kweon YO, Park JG, Lee HW, Han YS, Chun JM, Park SY et al: Circulating exosomal noncoding RNAs as prognostic biomarkers in human hepatocellular carcinoma. International journal of cancer 2019, 144(6).
15. Shukla S, Evans JR, Malik R, Feng FY, Dhanasekaran SM, Cao X, Chen G, Beer DG, Jiang H, Chinnaiyan AM: Development of a RNA-Seq Based Prognostic Signature in Lung Adenocarcinoma. J Natl Cancer Inst 2016, 109(1).
16. Li J, Wang W, Xia P, Wan L, Zhang L, Yu L, Wang L, Chen X, Xiao Y, Xu C: Identification of a five-lncRNA signature for predicting the risk of tumor recurrence in patients with breast cancer. International journal of cancer 2018, 143(9).
17. Qu L, Wang Z-L, Chen Q, Li Y-M, He H-W, Hsieh JJ, Xue S, Wu Z-J, Liu B, Tang H et al: Prognostic Value of a Long Non-coding RNA Signature in Localized Clear Cell Renal Cell Carcinoma. Eur Urol 2018, 74(6).
18. Zhou M, Hu L, Zhang Z, Wu N, Sun J, Su J: Recurrence-Associated Long Non-coding RNA Signature for Determining the Risk of Recurrence in Patients with Colon Cancer. Mol Ther Nucleic Acids 2018, 12
19. Luo X, Qiu Y, Jiang Y, Chen F, Jiang L, Zhou Y, Dan H, Zeng X, Lei YL, Chen Q: Long non-coding RNA implicated in the invasion and metastasis of head and neck cancer: possible function and mechanisms. Mol Cancer 2018, 17(1).
20. Pan J, Li X, Wu W, Xue M, Hou H, Zhai W, Chen W: Long non-coding RNA UCA1 promotes cisplatin/gemcitabine resistance through CREB modulating miR-196a-5p in bladder cancer cells. Cancer Lett 2016, 382(1).
21. Kia V, Mortazavi Y, Paryan M, Biglari A, Mohammadi-Yeganeh S: Exosomal miRNAs from highly metastatic cells can induce metastasis in non-metastatic cells. Life Sci 2019, 220
22. Liu H, Luo J, Luan S, He C, Li Z: Long non-coding RNAs involved in cancer metabolic reprogramming. Cell Mol Life Sci 2019, 76(3).
23. Schwarzenbacher D, Klec C, Pasculli B, Cerk S, Rinner B, Karbiener M, Ivan C, Barbano R, Ling H, Wulf-Goldenberg A et al: MiR-1287-5p inhibits triple negative breast cancer growth by interaction with

phosphoinositide 3-kinase CB, thereby sensitizing cells for PI3Kinase inhibitors. Breast Cancer Res 2019, 21(1).

24. Liu Z, Wang Y, Wang L, Yao B, Sun L, Liu R, Chen T, Niu Y, Tu K, Liu Q: Long non-coding RNA AGAP2-AS1, functioning as a competitive endogenous RNA, upregulates ANXA11 expression by sponging miR-16-5p and promotes proliferation and metastasis in hepatocellular carcinoma. J Exp Clin Cancer Res 2019, 38(1).
25. Zhou M, Guo M, He D, Wang X, Cui Y, Yang H, Hao D, Sun J: A potential signature of eight long non-coding RNAs predicts survival in patients with non-small cell lung cancer. J Transl Med 2015, 13
26. Zhu M, Lv Q, Huang H, Sun C, Pang D, Wu J: Identification of a four-long non-coding RNA signature in predicting breast cancer survival. Oncol Lett 2020, 19(1).
27. Fan Q, Liu B: Discovery of a novel six-long non-coding RNA signature predicting survival of colorectal cancer patients. J Cell Biochem 2018, 119(4).
28. Zhu X, Tian X, Yu C, Shen C, Yan T, Hong J, Wang Z, Fang JY, Chen H: A long non-coding RNA signature to improve prognosis prediction of gastric cancer. Mol Cancer 2016, 15(1).
29. He A, He S, Peng D, Zhan Y, Li Y, Chen Z, Gong Y, Li X, Zhou L: Prognostic value of long non-coding RNA signatures in bladder cancer. Aging (Albany NY) 2019, 11(16).
30. Sun Y, Zhang F, Wang L, Song X, Jing J, Zhang F, Yu S, Liu H: A five lncRNA signature for prognosis prediction in hepatocellular carcinoma. Mol Med Rep 2019, 19(6).
31. Li W, Chen QF, Huang T, Wu P, Shen L, Huang ZL: Identification and Validation of a Prognostic lncRNA Signature for Hepatocellular Carcinoma. Front Oncol 2020, 10
32. Ma L, Deng C: Identification of a novel four-lncRNA signature as a prognostic indicator in cirrhotic hepatocellular carcinoma. PeerJ 2019, 7
33. Zhao QJ, Zhang J, Xu L, Liu FF: Identification of a five-long non-coding RNA signature to improve the prognosis prediction for patients with hepatocellular carcinoma. World J Gastroenterol 2018, 24(30).
34. Huang JL, Zheng YS, Xiao XL, Liu C, Lin JP, Zheng SJ, Yang B, Ou QS: A Circulating Long Noncoding RNA Panel Serves as a Diagnostic Marker for Hepatocellular Carcinoma. Disease markers 2020, 2020
35. Yuan WD, Sun Y, Liu L, Zhou B, Wang SM, Gu DH: Circulating LncRNAs Serve as Diagnostic Markers for Hepatocellular Carcinoma. Cell Physiol Biochem 2017, 44(1).
36. Xiong D, Sheng Y, Ding S, Chen J, Tan X, Zeng T, Qin D, Zhu L, Huang A, Tang H: LINC00052 regulates the expression of NTRK3 by miR-128 and miR-485-3p to strengthen HCC cells invasion and migration. Oncotarget 2016, 7(30).
37. Ji J, Tang J, Deng L, Xie Y, Jiang R, Li G, Sun B: LINC00152 promotes proliferation in hepatocellular carcinoma by targeting EpCAM via the mTOR signaling pathway. Oncotarget 2015, 6(40).
38. Lee KK, Kim DG, Moon IS, Lee MD, Park JH: Liver transplantation versus liver resection for the treatment of hepatocellular carcinoma. J Surg Oncol 2010, 101(1).
39. Ho CM, Lee PH, Chen CL, Ho MC, Wu YM, Hu RH: Long-term outcomes after resection versus transplantation for hepatocellular carcinoma within UCSF criteria. Ann Surg Oncol 2012, 19(3).
40. Zheng Z, Liang W, Milgrom DP, Zheng Z, Schroder PM, Kong NS, Yang C, Guo Z, He X: Liver transplantation versus liver resection in the treatment of hepatocellular carcinoma: a meta-analysis of observational studies. Transplantation 2014, 97(2).

41. Gao J, Kwan PW, Shi D: Sparse kernel learning with LASSO and Bayesian inference algorithm. Neural Netw 2010, 23(2).

Tables

Table 1 A
The clinical characteristics between the high-risk and low-risk group divided by the DFS-related lncRNAs signature.

Variable	Training set			Validation set			Entire set		
	Number High/low	x-squared	p-value	Number High/low	x-squared	p-value	Number High/low	x-squared	p-value
Age		0.006	0.936		2.750	0.097		0.082	0.775
≤60	44/58			22/31			81/80		
>60	46/62			30/22			75/79		
pT		0.598	0.439		7.008	0.008		4.379	0.036
T1-2	65/90			35/47			100/137		
T3-4	25/27			17/6			42/33		
pN		0.057	0.811		1.943	0.163		1.508	0.219
N0	62/87			33/33			95/120		
N1	1/1			2/0			3/1		
pM		1.233	0.267		0.003	0.954		0.462	0.497
M0	70/87			38/35			108/122		
M1	1/0			1/1			2/1		
Stage		1.412	0.235		7.477	0.006		6.259	0.012
I-II	60/85			31/44			91/129		
III-IV	26/25			17/6			43/31		
Grade		9.463	0.002		8.353	0.004		17.708	<0.001
G1-2	47/86			25/40			72/126		
G3-4	42/31			27/13			69/44		

Table 1 B
The clinical characteristics between the high-risk and low-risk group divided by the OS-related lncRNAs signature.

Variable	Training set			Validation set			Entire set		
	Number High/low	x-squared	p-value	Number High/low	x-squared	p-value	Number High/low	x-squared	p-value
Age		0.067	0.796		0.447	0.504		0.346	0.556
≤60	53/64			28/28			81/92		
>60	55/71			29/37			84/108		
pT		7.326	0.007		2.279	0.131		9.279	0.002
T1-2	68/104			43/56			111/160		
T3-4	40/28			14/9			54/37		
pN		0.045	0.832		-	-		0.044	0.834
N0	71/88			40/49			111/137		
N1	2/2			0/0			2/2		
pM		0.856	0.355		1.794	0.180		2.610	0.106
M0	80/93			43/47			123/140		
M1	0/1			0/2			0/3		
Stage		6.195	0.013		2.565	0.109		8.166	0.004
I-II	62/98			42/52			104/150		
III-IV	37/28			14/8			51/36		
Grade		5.168	0.023		17.591	<0.001		17.800	<0.001
G1-2	57/90			28/55			85/145		
G3-4	49/42			29/10			78/52		

Table 2
Baseline data of HCC patients.

Characteristic	Classification	n (%)
Total		371
Age, years	>60	193 (52.0)
	≤60	177 (47.7)
	Not available	1 (0.3)
pT	T1	181 (48.8)
	T2	94 (25.3)
	T3	80 (21.5)
	T4	13 (3.5)
	Tx	1 (0.3)
	Not available	2 (0.6)
pN	N0	252 (67.9)
	N1	4 (1.1)
	Nx	114 (30.7)
	Not available	1 (0.3)
pM	M0	266 (71.7)
	M1	4 (1.1)
	Mx	101 (27.2)
Stage	Stage I	171 (46.1)
	Stage II	86 (23.2)
	Stage III	85 (22.9)
	Stage IV	5 (1.3)
	Not available	24 (6.5)
Grade	G1	55 (14.8)
	G2	177 (47.7)
	G3	122 (32.9)
	G4	12 (3.2)
	Not available	5 (1.3)

Figures

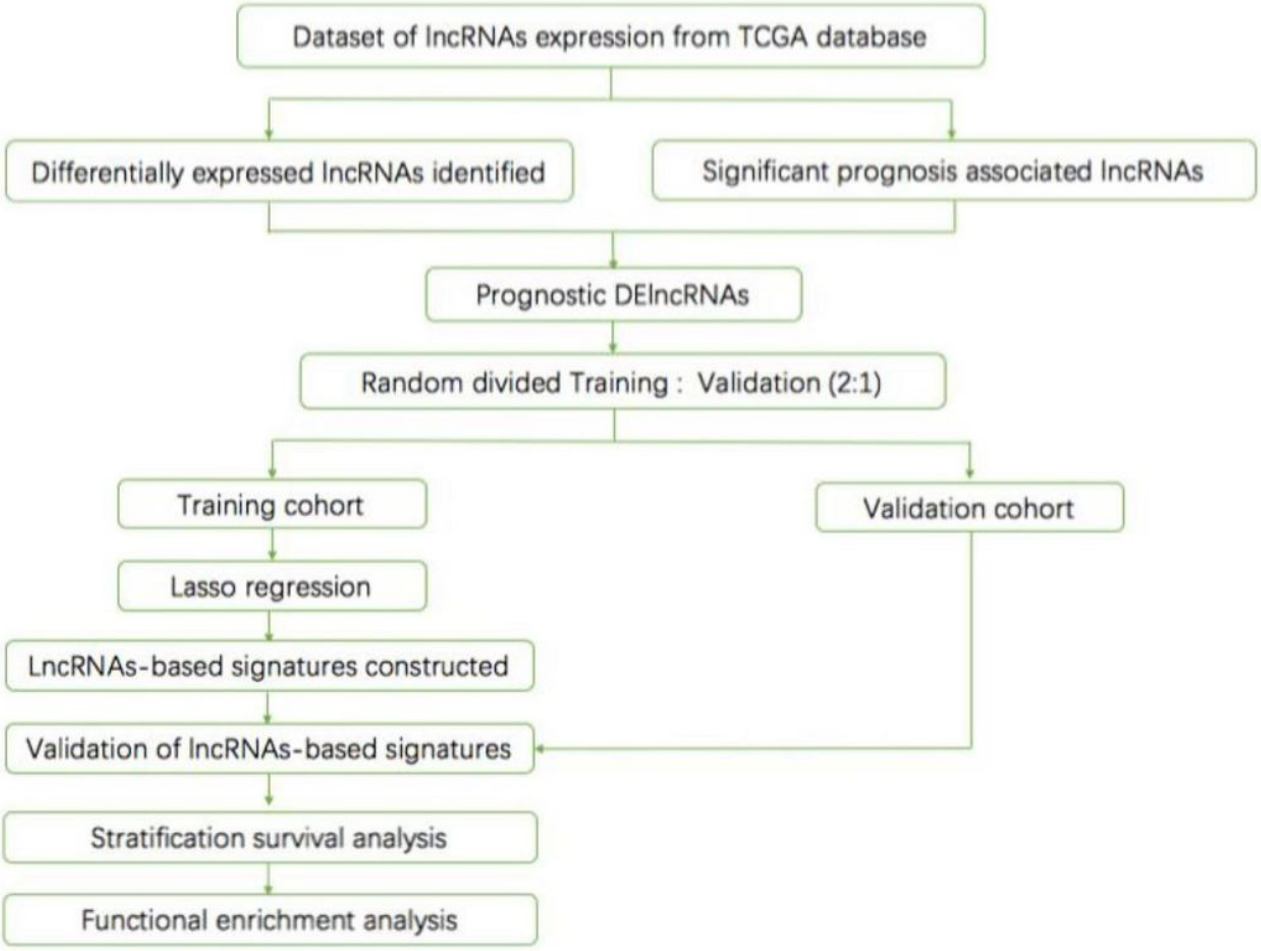

Figure 1

The workflow of this study.

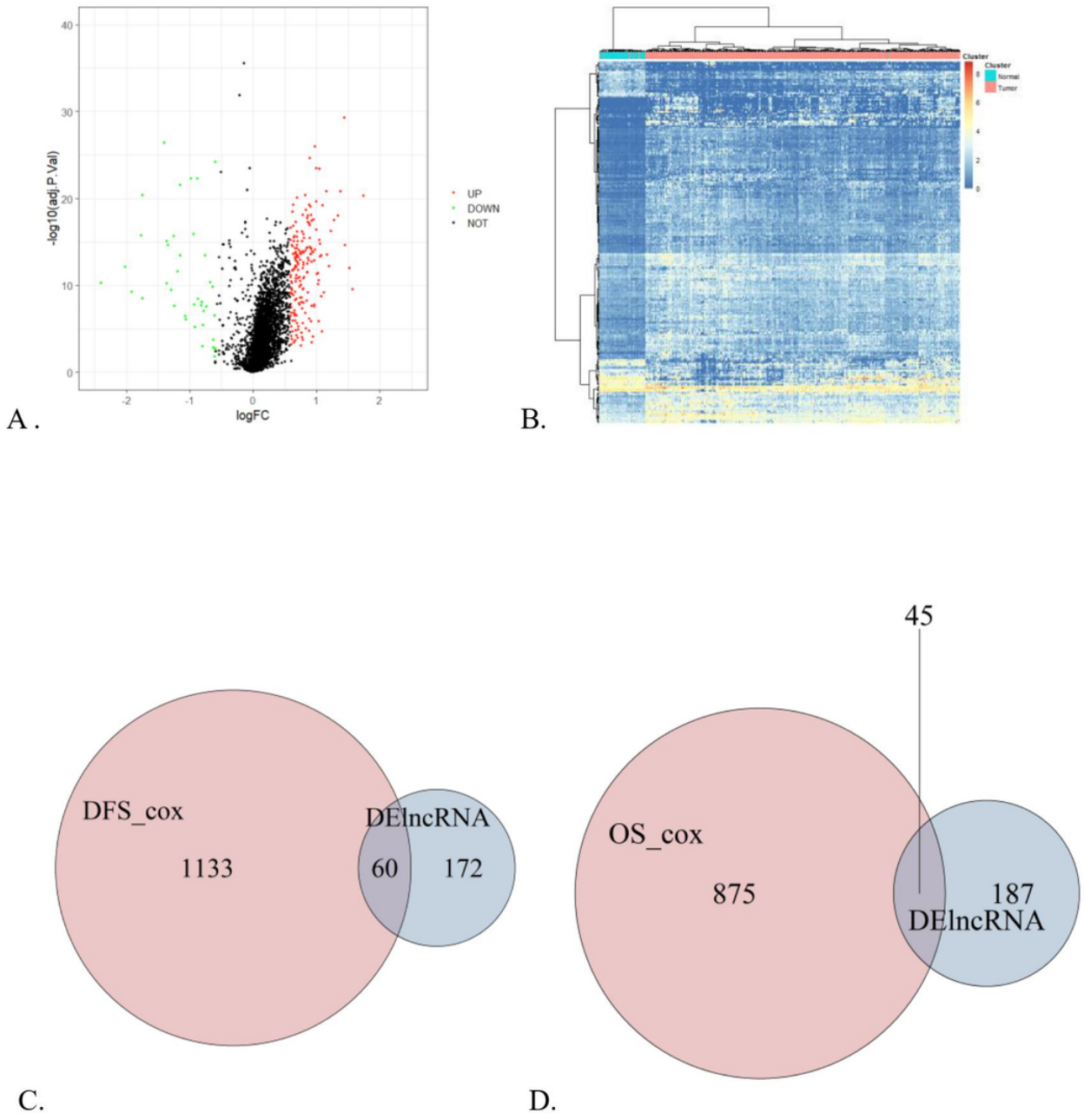

Figure 2

Prognosis-associated DElncRNAs analysis. (A) Volcano plot of DElncRNAs; (B) Heat map of DElncRNAs expressions; (C) DFS-related DElncRNAs selection; (D) OS-related DElncRNAs selection.

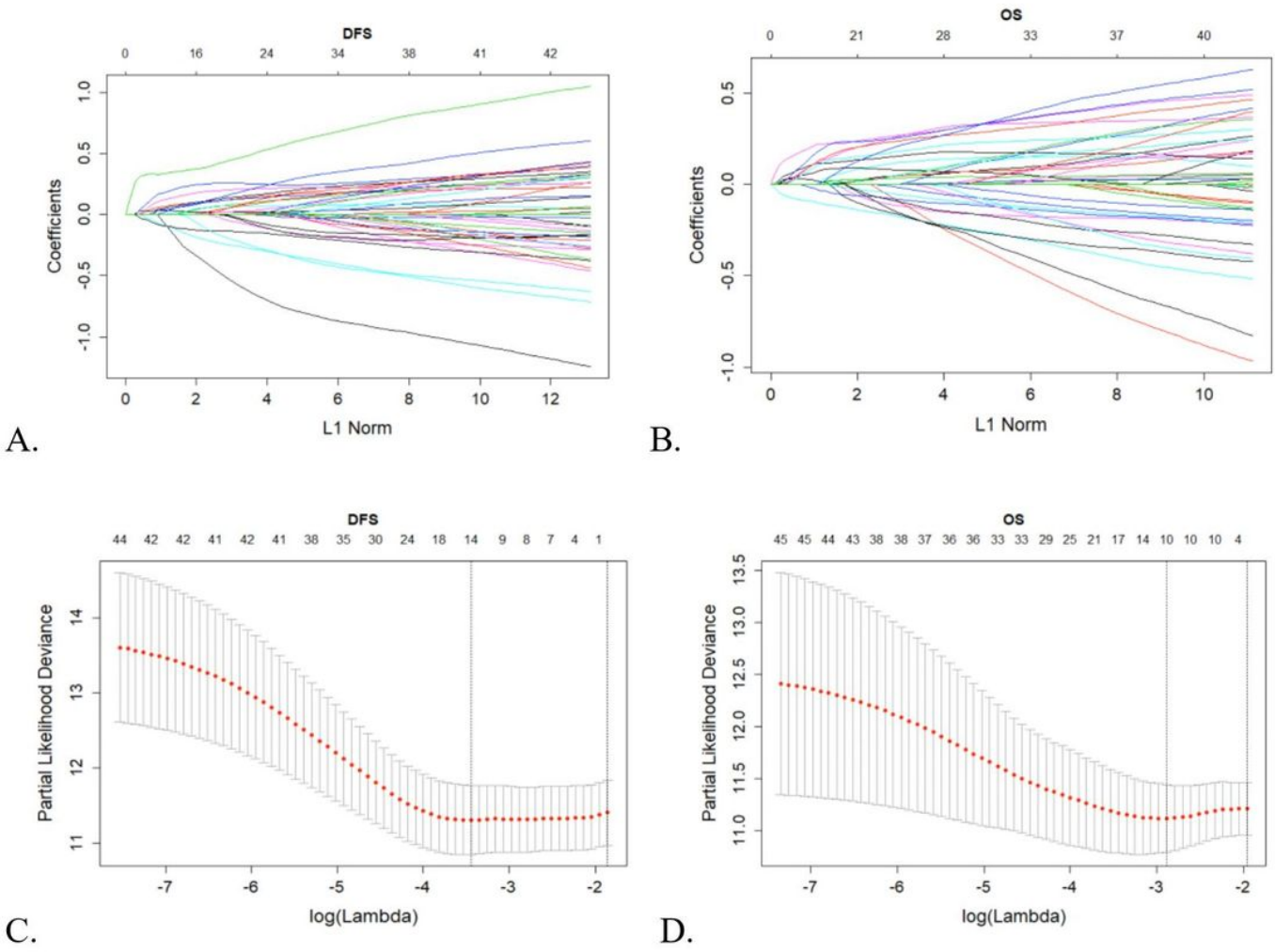

Figure 3

LASSO coefficient profiles of the significant DFS-related DElncRNAs (A) and OS-related DElncRNAs (B); Vertical lines were drawn with 20-fold cross validation for choosing the tunning parameters for DFS-related (C) and OS-related (D) LASSO models.

Figure 4

Kaplan-Meier analysis for high-risk group and low-risk group. The survival curve of DFS in training cohort (A), validation cohort (B) and entire cohort (C). The survival curve of OS in training cohort (D), validation cohort (E) and entire cohort (F).

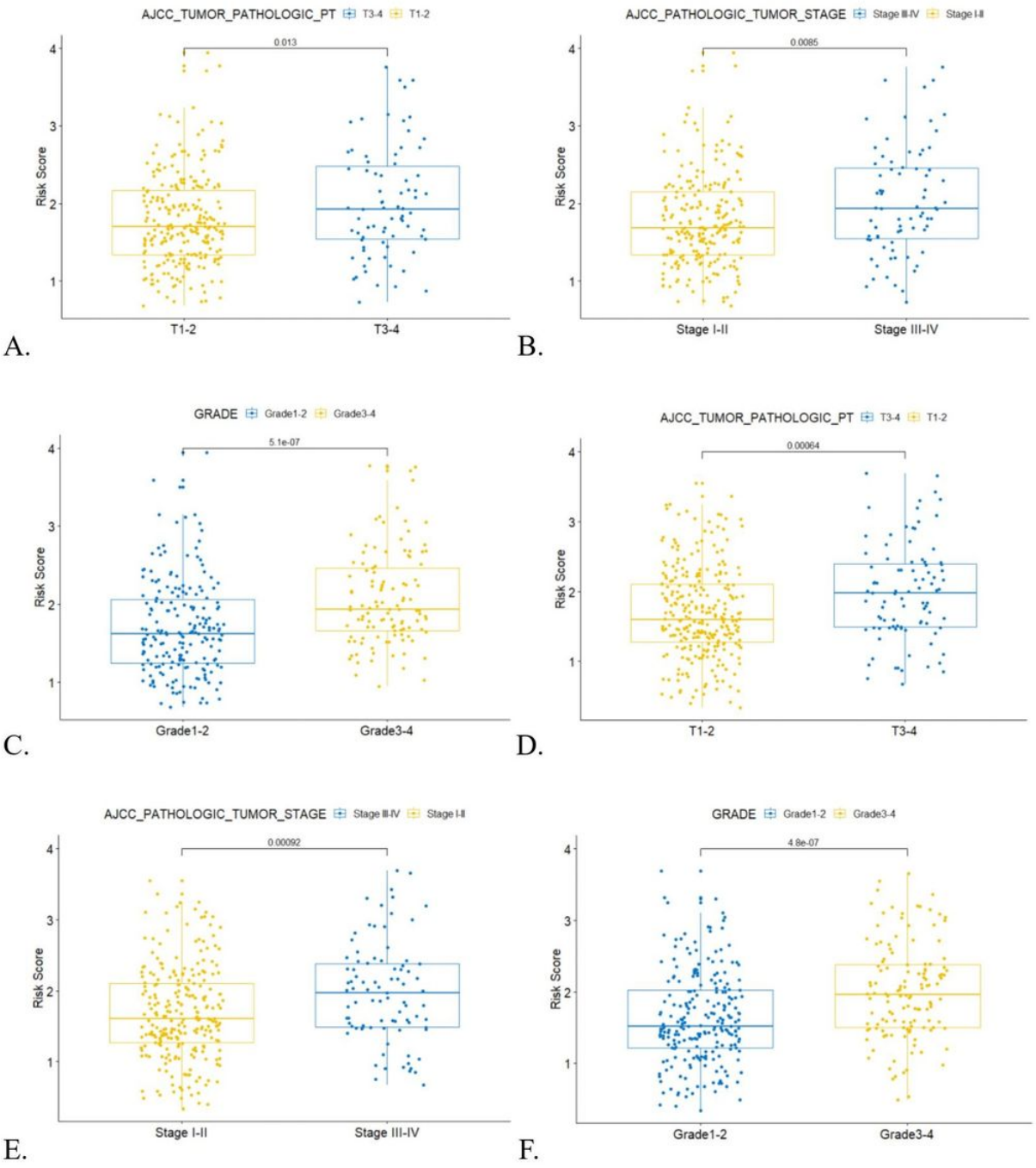

Figure 5

The correlation of risk score and clinicopathological characteristics in DFS and OS related cohort. (A)Risk score and pT in DFS related cohort. (B) Risk score and tumor stage in DFS related cohort. (C) Risk score and tumor grade in DFS related cohort. (D) Risk score and pT in OS related cohort. (E) Risk score and tumor stage in OS related cohort. (F) Risk score and tumor grade in OS related cohort.

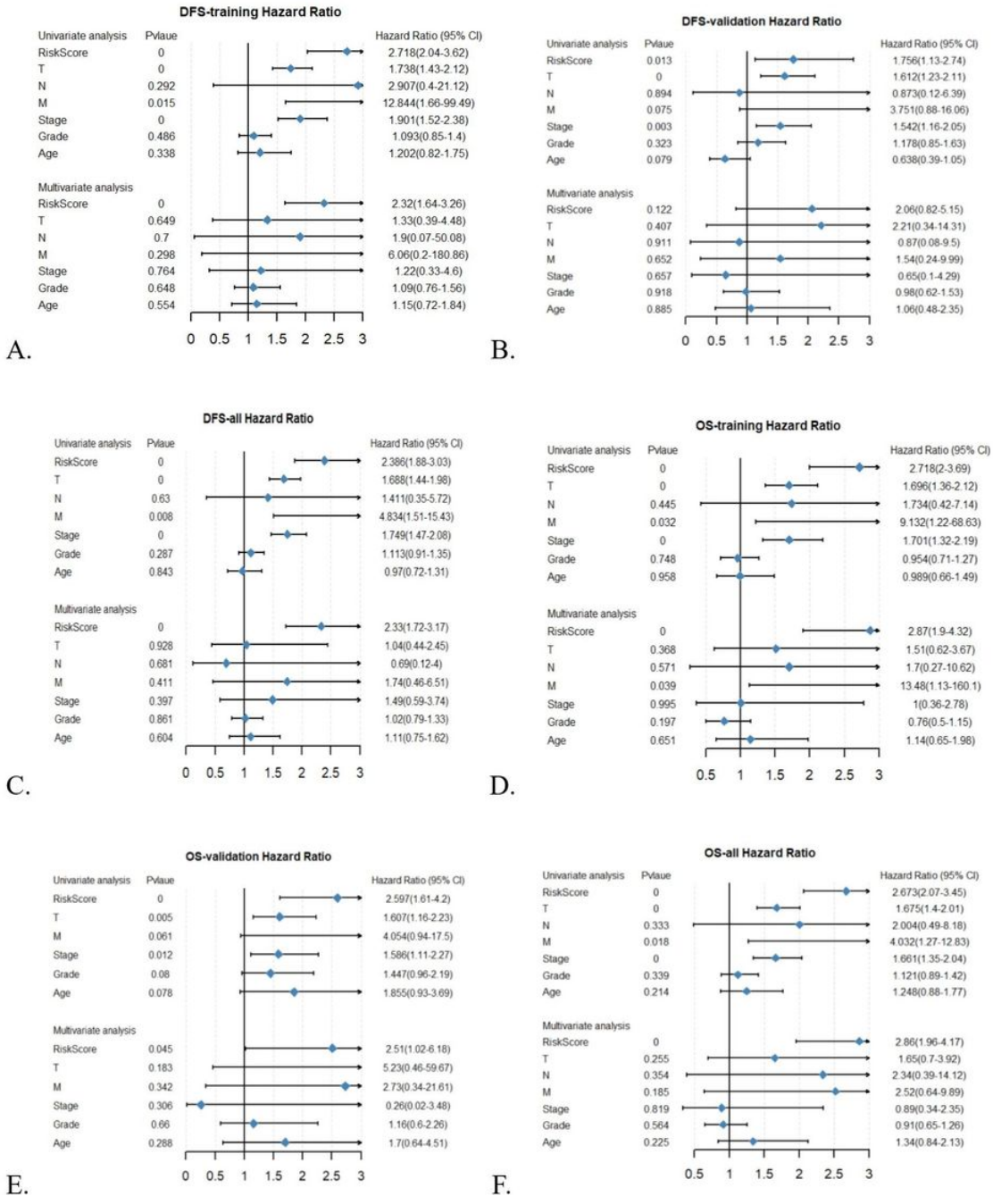

Figure 6

Cox regression results. Univariate and multivariate results in the DFS-related training (A), validation (B) and entire set (C). Univariate and multivariate results in the OS-related training (D), validation (E) and entire set (F).

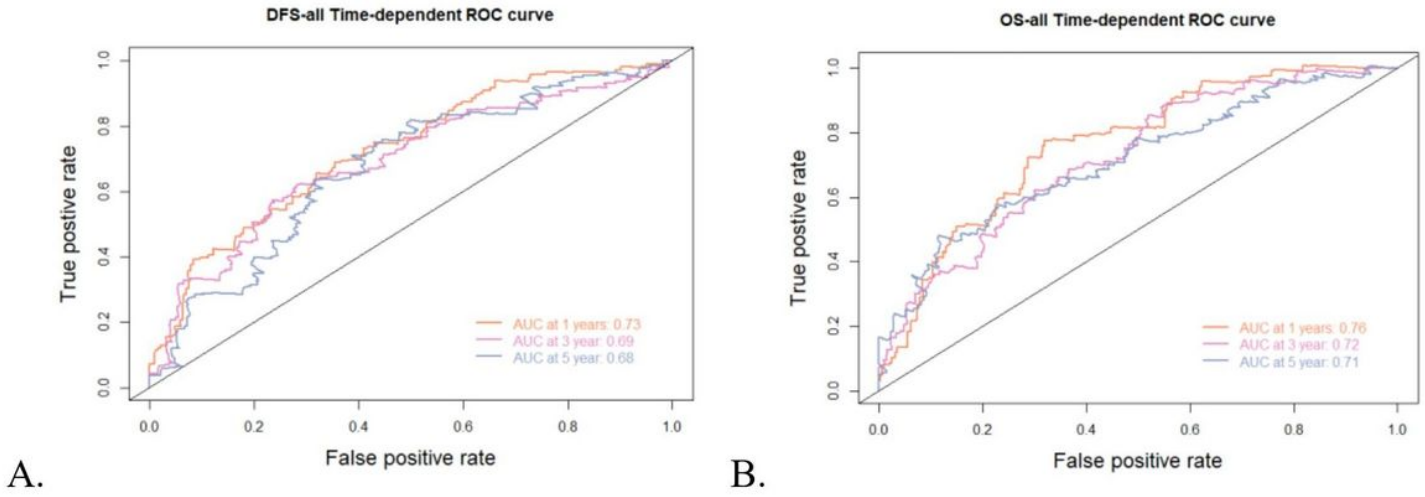

Figure 7

Receiver operating characteristics curves for DFS (A) and OS (B)in the entire set.

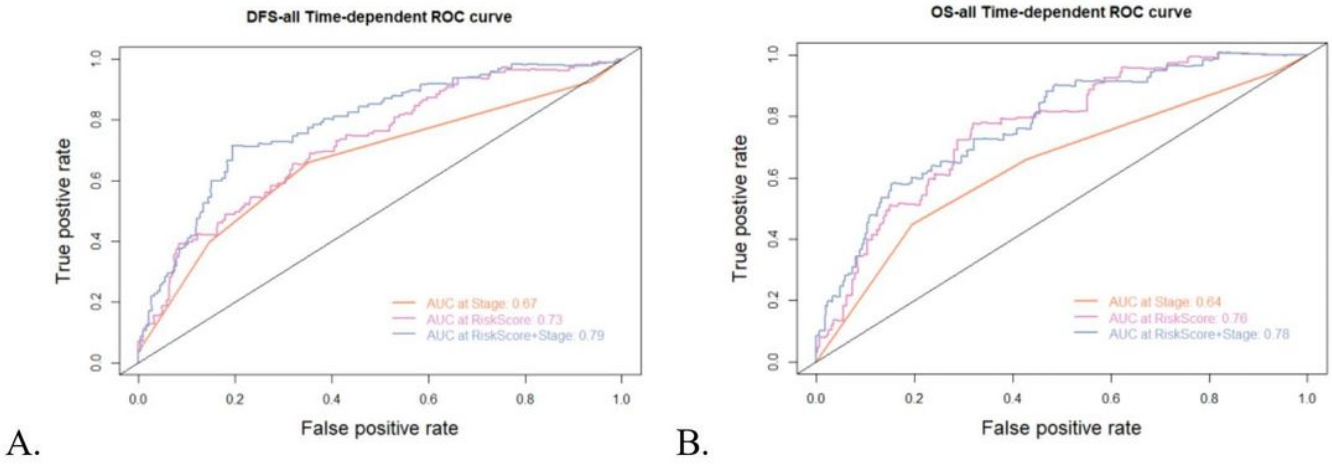

Figure 8

Receiver operating characteristics curves of lncRNA-based signature and AJCC 8th TNM stage for DFS(A) and OS (B)in the entire set.

Figure 9

Kaplan-Meier analysis for subgroup divided by combining stage of the tumor with risk score. The survival curve of DFS (A); the survival curve of OS (B).

Figure 10

Functional enrichment analysis of DFS-related lncRNAs: significantly enriched GO terms (A) and KEGG pathways (B); Functional enrichment analysis of OS-related lncRNAs: significantly enriched GO terms (C) and KEGG pathways (D).