



# A Long Non-coding RNA Signature to Improve Prognostic Prediction of Pancreatic Ductal Adenocarcinoma

## OPEN ACCESS

Chenhao Zhou<sup>1†</sup>, Shun Wang<sup>2†</sup>, Qiang Zhou<sup>1†</sup>, Jin Zhao<sup>2</sup>, Xianghou Xia<sup>3</sup>, Wanyong Chen<sup>1,4</sup>, Yan Zheng<sup>2</sup>, Min Xue<sup>4</sup>, Feng Yang<sup>5</sup>, Deliang Fu<sup>5</sup>, Yirui Yin<sup>1</sup>, Manar Atyah<sup>1</sup>, Lunxiu Qin<sup>2,4</sup>, Yue Zhao<sup>6</sup>, Christiane Bruns<sup>6</sup>, Huliang Jia<sup>2\*</sup>, Ning Ren<sup>1,7\*</sup> and Qiongzhu Dong<sup>2,4\*</sup>

### Edited by:

Zhaohui Huang,  
Affiliated Hospital of Jiangnan  
University, China

### Reviewed by:

Pieter J. Eichhorn,  
National University of  
Singapore, Singapore  
Suresh K. Kalangi,  
Amity University Gurgaon, India

### \*Correspondence:

Huliang Jia  
jbl-1@163.com  
Ning Ren  
renning@fudan.edu.cn  
Qiongzhu Dong  
qzhdong@fudan.edu.cn

<sup>†</sup>These authors have contributed  
equally to this work

### Specialty section:

This article was submitted to  
Cancer Molecular Targets and  
Therapeutics,  
a section of the journal  
Frontiers in Oncology

**Received:** 02 August 2019

**Accepted:** 17 October 2019

**Published:** 08 November 2019

### Citation:

Zhou C, Wang S, Zhou Q, Zhao J,  
Xia X, Chen W, Zheng Y, Xue M,  
Yang F, Fu D, Yin Y, Atyah M, Qin L,  
Zhao Y, Bruns C, Jia H, Ren N and  
Dong Q (2019) A Long Non-coding  
RNA Signature to Improve Prognostic  
Prediction of Pancreatic Ductal  
Adenocarcinoma.  
*Front. Oncol.* 9:1160.  
doi: 10.3389/fonc.2019.01160

<sup>1</sup> Department of Liver Surgery, Liver Cancer Institute, Zhongshan Hospital, Fudan University, Shanghai, China, <sup>2</sup> Department of General Surgery, Huashan Hospital & Cancer Metastasis Institute, Fudan University, Shanghai, China, <sup>3</sup> Department of Breast Surgery, Zhejiang Cancer Hospital, Zhejiang, China, <sup>4</sup> Institute of Biomedical Sciences, Fudan University, Shanghai, China, <sup>5</sup> Department of Pancreatic Surgery, Pancreatic Disease Institute, Huashan Hospital, Fudan University, Shanghai, China, <sup>6</sup> Department of General, Visceral and Cancer Surgery, University Hospital of Cologne, Cologne, Germany, <sup>7</sup> Institute of Fudan Minhang Academic Health System, Minhang Hospital, Fudan University, Shanghai, China

**Background:** Pancreatic ductal adenocarcinoma (PDAC) remains one of the most aggressive solid malignant tumors worldwide. Increasing investigations demonstrate that long non-coding RNAs (lncRNAs) expression is abnormally dysregulated in cancers. It is crucial to identify and predict the prognosis of patients for the selection of further therapeutic treatment.

**Methods:** PDAC lncRNAs abundance profiles were used to establish a signature that could better predict the prognosis of PDAC patients. The least absolute shrinkage and selection operator (LASSO) Cox regression model was applied to establish a multi-lncRNA signature in the TCGA training cohort ( $N = 107$ ). The signature was then validated in a TCGA validation cohort ( $N = 70$ ) and another independent Fudan cohort ( $N = 46$ ).

**Results:** A five-lncRNA signature was constructed and it was significantly related to the overall survival (OS), either in the training or validation cohorts. Through the subgroup and Cox regression analyses, the signature was proven to be independent of other clinic-pathologic parameters. Receiver operating characteristic curve (ROC) analysis also indicated that our signature had a better predictive capacity of PDAC prognosis. Furthermore, ClueGO and CluePedia analyses showed that a number of cancer-related and drug response pathways were enriched in high risk groups.

**Conclusions:** Identifying the five-lncRNA signature (RP11-159F24.5, RP11-744N12.2, RP11-388M20.1, RP11-356C4.5, CTC-459F4.9) may provide insight into personalized prognosis prediction and new therapies for PDAC patients.

**Keywords:** pancreatic ductal adenocarcinoma, lncRNA, signature, prognosis, overall survival

## INTRODUCTION

As reported in Cancer Statistics (2019), pancreatic cancer (PC), one of the most aggressive solid malignant tumors, is the fourth cause of cancer related deaths in the USA and it is estimated that 56,770 PC cases occurred and 45,750 patients died from PC in 2019. Additionally, PC incidence rate and death rate continues to increase, and a 5-year relative survival rate for it is the lowest (9%) compared with other cancers (1). Furthermore, pancreatic ductal adenocarcinoma (PDAC) accounts for the most part of PC (2). Due to atypical symptoms of PDAC and ineffective diagnosis methods, the majority of PDAC patients were diagnosed at an advanced stage, leading to patients not having the chance to be resected (1–4). Given current medical capabilities, surgical resection for resectable PDAC patients offers the only chance of a cure, but the survival of patients has not improved and commonly recurs within 12 months (5, 6). As for PDAC patients being diagnosed at an advanced stage, chemotherapy such as nab-paclitaxel plus gemcitabine and other methods including radiotherapy, molecular targeted therapy (MTT), and immunotherapy have only yielded modest improvements in survival (7, 8). Therefore, further exploration concerning the molecular mechanism underlying PDAC occurrence and progression are essential to improve the early diagnostic biomarkers and therapeutic targets.

Long non-coding RNAs (lncRNAs) are defined as transcripts with a length of over 200 nucleotides, and lack protein-coding potential (9, 10). Studies have demonstrated that lncRNAs play important roles in a variety of biological processes, including epigenetic regulation, alternative splicing, imprinting, cell cycle control, cell differentiation, drug resistance, and tumorigenesis (11–14). Furthermore, emerging investigations have indicated that lncRNAs expression is abnormally dysregulated in numerous cancers and many lncRNAs are correlated with cancer recurrence, metastasis, and poor prognosis (15–17). For instance, the dysregulation of lncRNAs, including Linc01060, AGAP2-AS1, LINC00958, participate in the incidence, metastasis and progression of PDAC (18–20). Carbohydrate antigen 19-9 (CA19-9) is currently confirmed as the prognostic biomarkers for PDAC, but its low specificity urges us to discover and identify more potential and valuable molecular biomarkers for patients with PDAC (21). Increasing evidence suggests that lncRNA could serve as a good choice in predicting the prognosis of cancer (22–24). Meanwhile, a lot of gene expression signatures were successfully established to predict the clinical outcome of many different types of cancer (25, 26).

There are many lncRNAs which have reportedly been associated with PDAC prognosis (27, 28). However, they are seldomly used in clinical practice considering that their signature

has not been validated in an independent cohort, or they do not adopt the appropriate statistical approach to generate the model. Here, we aim to identify the potential minimum number of robust lncRNAs as a signature to predict the prognosis of PDAC patients. Therefore, we mined the lncRNA data from The Cancer Genome Atlas (TCGA) database using the least absolute shrinkage and selection operator method (LASSO) algorithm, which can effectively analyze the high-dimensional sequencing data (29). We then validated this 5-lncRNA signature in our own PDAC patients from Fudan University using qRT-PCR technology and evaluated the accuracy of this signature and predictive capacity in the entire TCGA cohort.

## METHODS

### PDAC Datasets Preparation

The level three RNA sequencing data and relevant clinical information of 177 PDAC patients were downloaded from the TCGA database (<http://cancergenome.nih.gov/>) and enrolled in our study. Additionally, 46 fresh frozen primary PDAC samples from the Fudan validation cohort were collected consecutively at Huashan Hospital from October 2010 to February 2014. All enrolled patients met the inclusion and exclusion criteria as follows: (1) pathologic diagnosis of PDAC without other types of pancreatic cancer; (2) no cases with other malignant cancers. Conventional clinicopathologic variables like age, gender, AJCC Tumor Node Metastasis (TNM) stage, histologic grade, and microsatellite instability (MSI) status were analyzed in our study. Written informed consent was obtained from all patients. The study was conducted in accordance with the Declaration of Helsinki, and the Ethical Committee of Huashan Hospital, Fudan University approved the study.

### RNA-seq Data Processing and lncRNA Profile Mining

To obtain the lncRNA expression profile, we mapped the probes of the TCGA RNA-seq data to lncRNA annotation files in the GENCODE database (<http://www.genecodegenes.org>). The annotation data (antisense, lincRNA, and sense\_intronic) of probes was recognized as lncRNA. Fifteen thousand eight hundred and ninety-nine lncRNA probes were acquired in the RNA-seq data of PDAC. After removing lncRNAs whose expression was zero in more than 20% of the samples, 6,010 annotated lncRNA probes were generated for further study.

### RNA Extraction and Quantitative Reverse Transcription PCR (qRT-PCR)

TRIzol reagent (Invitrogen, USA) was used to extract total RNA from 46 PDAC samples (Fudan validation cohort) following the manufacturer's protocol. Reverse-transcription was then carried out using a PrimeScript RT reagent kit (Takara, Japan). An ABI Prism 7500 Sequence Detection System (Applied Biosystems, Foster City, CA, USA) was used to carry out the qRT-PCR using SYBR® Premix ExTaq™ (Takara, Japan). ACTB ( $\beta$ -actin) was utilized as an internal control to normalize the expression of lncRNAs. The  $-\Delta\text{CT}$  method ( $\Delta\text{CT} = \text{CT lncRNA} - \text{CT ACTB}$ )

**Abbreviations:** AUROC, Area under receiving operating curve; CI, Confidence interval; DEGs, Differentially expressed genes; FC, Fold change; HR, Hazard ratio; LASSO, least absolute shrinkage and selection operator; lincRNA, large intergenic non-coding RNAs; lncRNAs, long non-coding RNAs; MSI, microsatellite instability; MSI-I, MSI-indeterminate; MSI-L, MSI-low; MSS, microsatellite stable; OS, Overall survival; PC, pancreatic cancer; PDAC, Pancreatic ductal adenocarcinoma; ROC, Receiver operating characteristic curve; TCGA, The Cancer Genome Atlas; TNM, Tumor Node Metastasis.

RNA) was applied to calculate each lncRNA expression level. The primers of related lncRNAs are shown in **Table S1**.

## Identification and Validation of the Prognostic lncRNA Signature

First, all the TCGA PDAC patients were randomly divided into two cohorts: the first 107 patients (60%) were termed the training cohort, and the remaining 70 (40%) the validation cohort. A univariate Cox regression model was applied to the training cohort to detect the prognostic lncRNAs. A set of lncRNAs whose *P*-value was <0.05 were identified. These lncRNAs were then analyzed in a training cohort utilizing R software (version 3.6.0) and the “glmnet” package (R Foundation for Statistical Computing, Vienna, Austria) to carry out the LASSO Cox regression model analysis. Ten-times cross-validations were used to find the best penalty parameter lambda (29, 30). A list of prognostic lncRNAs with related coefficients were obtained from the lncRNA expression profile and the patient's overall survival (OS) according to the best lambda value. Furthermore, the risk score of every patient was calculated according to the expression level of each prognostic lncRNA and its corresponding coefficient. PDAC patients were assigned to a high-risk or low-risk group using the median risk score as the cut-off point. In the end, the OS differences between the two groups were evaluated by the Kaplan–Meier survival curves. Meanwhile, the prognostic lncRNA signature was validated in the TCGA and Fudan validation cohort. Based on the median risk score, the validation cohort was also split into a high-risk or low-risk group, and OS differences were analyzed as described earlier.

## Statistical Analysis

All statistical analyses were performed using R software (version 3.6.0) and Bioconductor (31). For further analysis, the lncRNA expression profile was log<sub>2</sub>-transformed. Univariable, multivariable Cox regression, and stratified analyses were performed to evaluate if the lncRNA signature was independent of age, gender, TNM stage, grade, and MSI status. Student's *t* or the Fisher's exact test was used to assess the relationship between lncRNA signature and other clinicopathologic variables. The method of Kaplan–Meier and the log-rank test were utilized to evaluate the OS differences between the high or low-risk groups. Receiver operating characteristic (ROC) analysis was further used to assess the prognostic value based on the multi-lncRNA risk score, histologic grade, TNM stage, and the combined model of risk score and other indexes. “pROC” package was applied to the ROC curve analysis, and “delong” methodology was adopted to study significant differences between the ROC curves. When a two-sided *P*-value was <0.05, the statistical analyses were defined as statistically significant.

## Functional Enrichment Analysis

Differentially expressed genes (DEGs) between the high and low-risk PDAC patients in the TCGA dataset were identified using the classical *t*-test. The top 1,000 up-regulated and down-regulated DEGs were included in the functional enrichment analysis (**Tables S2, S3**). The Database for Annotation, Visualization and Integrated Discovery (DAVID, <https://david.ncicrf.gov/>) and Cytoscape plug-in ClueGO and CluePedia were used to perform the enrichment analysis (32–34). The threshold for the analyses was set as a *P* < 0.05. By using Cytoscape software, the significant functional enrichment results were visualized in our study.

ncicrf.gov/) and Cytoscape plug-in ClueGO and CluePedia were used to perform the enrichment analysis (32–34). The threshold for the analyses was set as a *P* < 0.05. By using Cytoscape software, the significant functional enrichment results were visualized in our study.

## RESULTS

### Clinical Characteristics of PDAC Patients

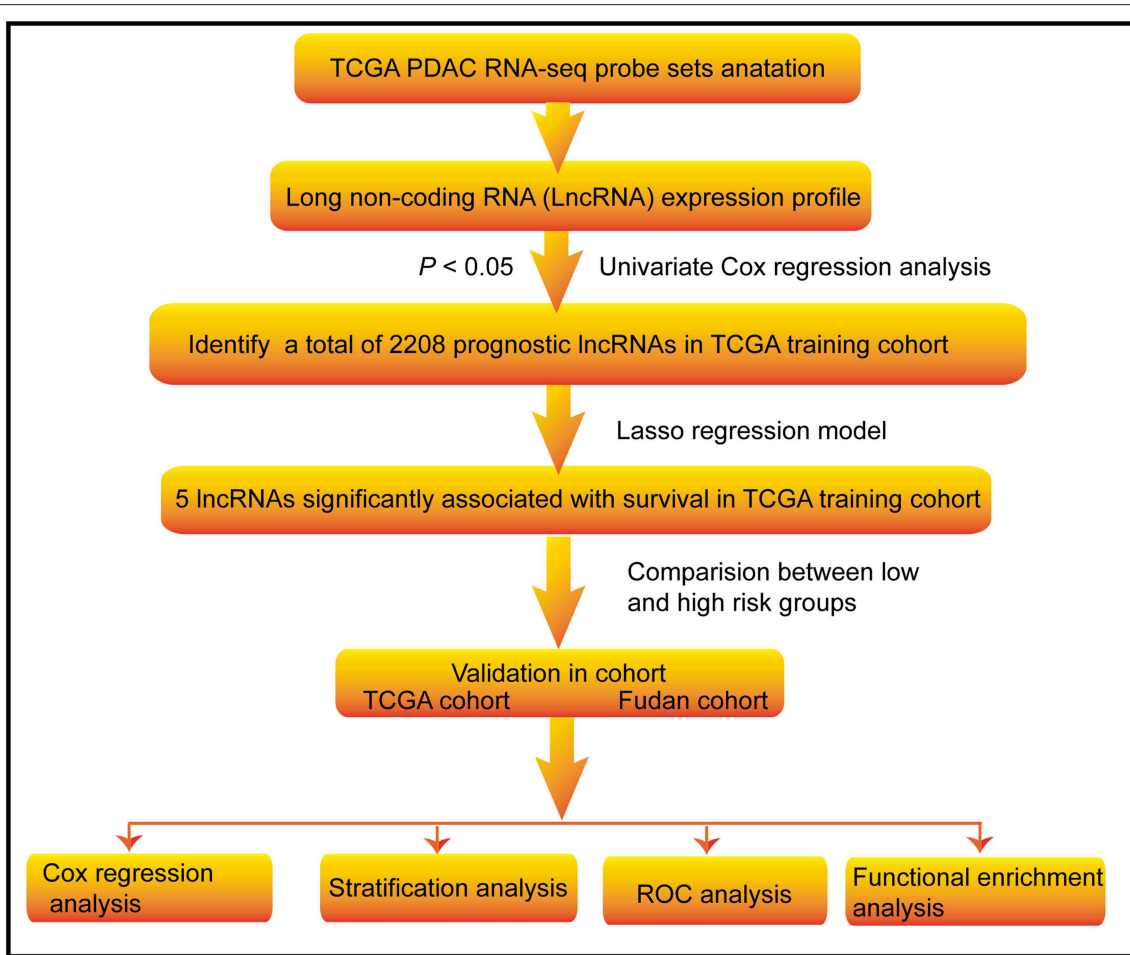
The flowchart of this study is shown in **Figure 1**. A total of 223 PDAC patients were enrolled in our study. The detailed clinical characteristics of these patients are summarized in **Table 1**. The 177 TCGA PDAC patients were randomly assigned to a TCGA training cohort (*N* = 107) and TCGA validation cohort (*N* = 70). Additionally, 46 PDAC patients from Huashan Hospital, Fudan University were recruited as another independent validation cohort. As shown in **Table 1**, the vast majority of patients (65.47%) were aged over 60 years, and about 54.71% of PDAC patients were male. One-hundred-and-forty TCGA patients (79.10%) has a microsatellite stable (MSS) status. Moreover, most tumors were diagnosed in well-differentiated groups (histologic grade I&II; 71.43%) and in early stage groups (TNM I&II; 90.37%).

### Generate Prognostic lncRNAs From TCGA Training Cohort

Using the univariate Cox regression analysis method, a set of 2,208 prognostic lncRNAs were identified in the TCGA training cohort (*P* < 0.05). A LASSO Cox regression model was further applied to those 2,208 lncRNAs to generate a prognostic signature in the training cohort. As a result, we recognized the five-lncRNA signature that was highly associated with OS in PDAC patients. A list of lncRNAs with their acquired coefficients, gene symbol, ensembl ID, gene type, *P*-values, and hazard ratio are demonstrated in **Table 2**. Among those lncRNAs, lower lncRNA expression levels were revealed by negative coefficients. Interestingly, the five lncRNAs identified, had completely negative coefficients - RP11-159F24.5, RP11-744N12.2, RP11-388M20.1, RP11-356C4.5, and CTC-459F4.9, which meant they were correlated with better survival.

### The 5-lncRNA Signature and the Patients' Survival in the Training Cohort

Based on the expression of these five lncRNAs for OS prediction, we established a risk-score formula: Risk score = (−0.00507\*expression level of RP11-159F24.5) + (−0.019766164\* expression level of RP11-744N12.2) + (−0.003146176\* expression level of RP11-388M20.1) + (−0.046208838\* expression level of RP11-356C4.5) + (−0.037384417\* expression level of CTC-459F4.9). Furthermore, we worked out the 5-lncRNA signature risk score for every patient in the TCGA training cohort. Using the median risk score as the cut-off point, the patients were categorized into a low risk group (*N* = 54) and high-risk group (*N* = 53). The Kaplan–Meier survival analysis indicated that patients in the high-risk group were



**FIGURE 1** | Flow chart of the study. The study was carried out in TCGA and Fudan lncRNA dataset of PDAC patients. The TCGA training cohort was used to identify prognostic lncRNAs. The LASSO regression model was used to establish a prognostic signature based on the prognostic lncRNAs. The prognosis analysis was validated in the TCGA and Fudan validation cohort, respectively.

associated with a tendency toward worse outcomes in the training cohort (HR = 2.03, 95% CI = 1.15–3.57,  $P = 0.012$ ) (Figure 2A).

### Validation of the 5-lncRNA Signature for Survival Prediction in the Validation Cohorts and Entire TCGA Cohort

In order to confirm the power of the 5-lncRNA signature in predicting the OS of PDAC patients, we validated our results in the internal validation cohort. By utilizing the same classification method, patients were classified into a high-risk group ( $N = 35$ ) and a low risk group ( $N = 35$ ). Consistent with previous findings, patients in the high-risk group revealed significantly worse OS compared to the other patients (HR = 6.77, 95% CI = 3.03–15.11,  $P < 0.0001$ ) (Figure 2B). Moreover, in the entire TCGA cohort, which included the training and validation cohort, the 5-lncRNA signature also had similar results (HR = 3.13, 95% CI = 1.99–4.93,  $P < 0.0001$ ) (Figure 2C). What is more, we validated the risk score-based classification in the independent cohort

from Huashan hospital, Fudan university. The Fudan validation cohort verified the capacity of the signature in predicting OS. As shown in Figure 2D, our signature can also effectively discriminate the risk of OS (HR = 3.25, 95% CI = 1.33–7.94,  $P = 0.0062$ ).

### Association Between 5-lncRNA Signature and Clinic-Pathologic Characteristics in PDAC Patients

The patients of the TCGA cohort and the Fudan validation cohort were divided into two groups for the purpose of investigating the importance of the 5-lncRNA signature in PDAC clinic-pathologic sides. As depicted in Table S4, a high-risk score of the signature was associated with aggressive tumor clinic-pathologic parameters, including TNM stage ( $P = 0.024$ , TCGA cohort) and OS status ( $P < 0.001$ , TCGA cohort;  $P = 0.008$ , Fudan validation cohort). Although there was no statistical significance of the histologic grade in the TCGA cohort and TNM stage in the Fudan validation cohort, it still had the trend that

**TABLE 1** | Clinical characteristics of 223 pancreatic adenocarcinoma patients involved in the study.

Characteristics	All (N = 223)	Detailed data		
		Training cohort (N = 107)	Validation cohort (N = 70)	Fudan cohort (N = 46)
Age at diagnosis, years				
≤60	77 (34.53%)	35 (15.70%)	19 (8.52%)	23 (10.31%)
≥60	146 (65.47%)	72 (32.29%)	51 (22.87%)	23 (10.31%)
Gender				
Female	101 (45.29%)	50 (22.42%)	30 (13.45%)	21 (9.42%)
Male	122 (54.71%)	57 (25.56%)	40 (17.94%)	25 (11.21%)
MSI status				
MSI-I	28 (15.82%)	14 (7.91%)	14 (7.91%)	–
MSI-L	9 (5.08%)	2 (1.13%)	7 (3.95%)	–
MSS	140 (79.10%)	91 (51.41%)	49 (27.68%)	–
Histologic grade				
G1	30 (17.14%)	22 (12.57%)	8 (4.57%)	–
G2	95 (54.29%)	55 (31.43%)	40 (22.86%)	–
G3	48 (27.43%)	28 (16%)	20 (11.43%)	–
G4	2 (1.14%)	1 (0.57%)	1 (0.57%)	–
TNM stage				
I	31 (14.22%)	12 (5.50%)	9 (4.13%)	10 (4.59%)
II	166 (76.15%)	88 (40.37%)	57 (26.15%)	21 (9.63%)
III	15 (6.88%)	2 (0.92%)	2 (0.92%)	11 (5.05%)
IV	6 (2.75%)	4 (1.83%)	1 (0.46%)	1 (0.46%)

MSI, Microsatellite instability; MSS, Microsatellite stable; MSI-I, MSI-indeterminate; MSI-L, MSI-low; TNM, Tumor node metastasis.

**TABLE 2** | lncRNAs significantly associated with the overall survival.

Gene symbol	Ensembl ID	Coefficient	Gene type	P-value	Hazard ratio
RP11-159F24.5	ENSG00000248240.1	−0.00507	Antisense	0.001618	0.608225
RP11-744N12.2	ENSG00000254703.2	−0.01977	Antisense	0.03337	0.756667
RP11-388M20.1	ENSG00000260060.1	−0.00315	Antisense	0.000288	0.518596
RP11-356C4.5	ENSG00000261172.1	−0.04621	LincRNA	0.013033	0.724935
CTC-459F4.9	ENSG00000281468.1	−0.03738	Sense_intronic	3.08E-05	0.568016

patients in the high-risk group showed a poor differentiated grade and an advanced tumor stage.

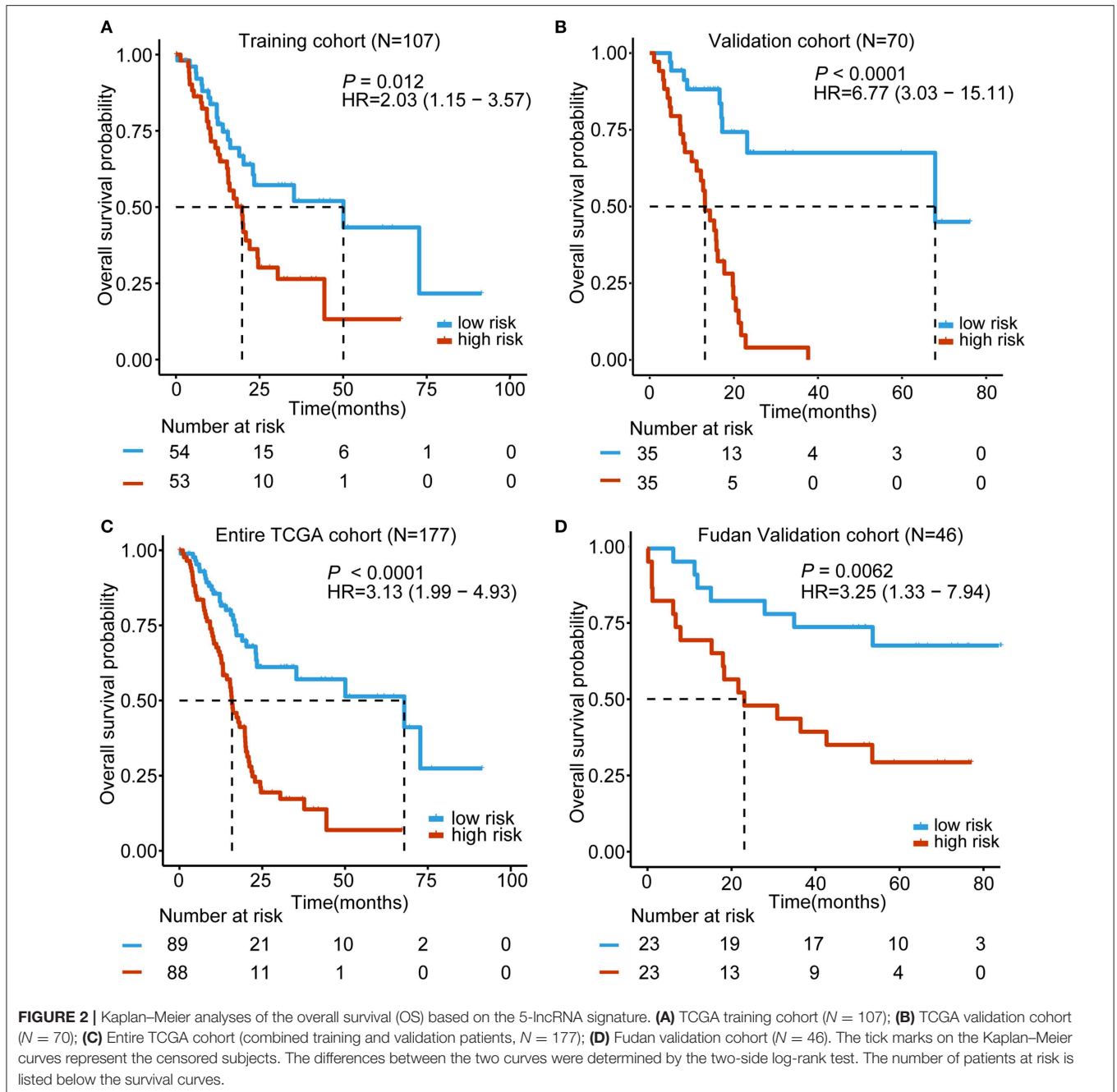
## Investigate the 5-lncRNA Signature Prognostic Capacities by Univariate and Multivariate Analyses

To confirm whether the prognostic capacity of our 5-lncRNA signature was independent from the clinic-pathologic characteristics, univariate and multivariate Cox analyses were carried out by analyzing the available co-variables like 5-lncRNA risk score, age, gender, TNM stage, histologic grade, and MSI status in the entire TCGA cohort and Fudan validation cohort. In the univariate Cox regression analyses, the 5-lncRNA signature was a powerful variable associated with prognosis in both the entire TCGA and Fudan validation cohort (HR = 3.10, 95% CI = 2.00–4.90,  $P < 0.0001$ ; HR = 3.25, 95% CI = 1.33–7.93,  $P = 0.01$ , respectively) (Figures 3A,C). By using other clinical

variables to adjust the multivariate analyses, the 5-lncRNA signature still proved to be a strong and independent variable in the above described cohorts (HR = 2.87, 95% CI = 1.82–4.51,  $P < 0.0001$ ; HR = 2.85, 95% CI = 1.12–7.27,  $P = 0.028$ , respectively) (Figures 3B,D).

## Subgroup Analyses Based on the 5-lncRNA Signature According to TNM Stage, Histologic Grade, and MSI Status

For the purpose of testing whether our 5-lncRNA signature can play a role in different TNM stages, histologic grades and MSI status, subgroup analyses were carried out, respectively. Because the number of TNM stage III&IV and not MSS status patients was small, we performed our subgroup analysis in the early stage (TNM stage I&II) and MSS status of TCGA PDAC patients. Figure 4A shows that our 5-lncRNA signature could successfully predict the survival outcome in this subgroup

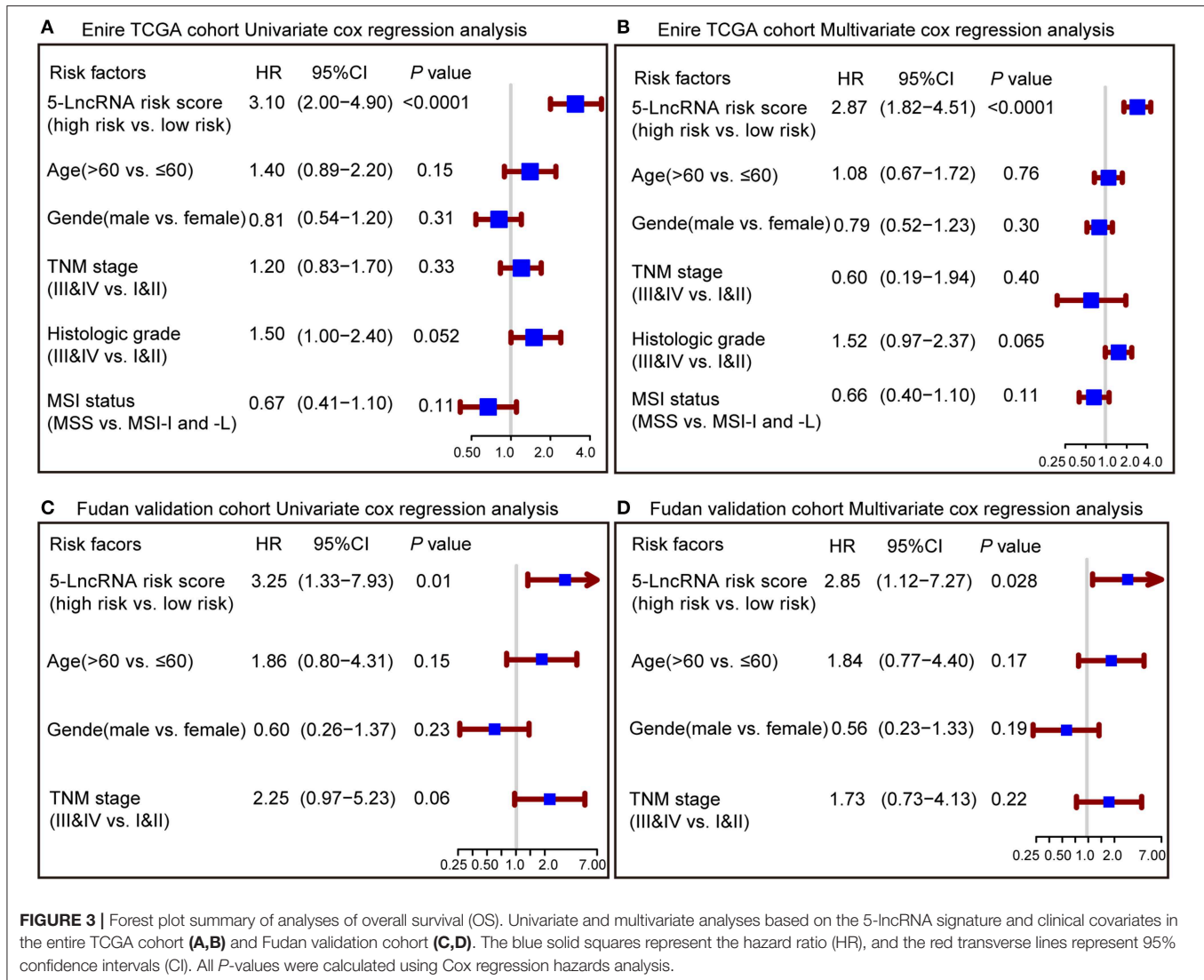


( $HR = 2.83$ ,  $95\% \text{ CI} = 1.79\text{--}4.46$ ,  $P < 0.0001$ ). Furthermore, TCGA PDAC patients were stratified into a well-differentiated group (histologic grade I&II) and poor-differentiated group (histologic grade III&IV). The subgroup analyses demonstrated that the 5-lncRNA signature could be beneficial to divide patients into low risk and high risk groups in every grade with a statistically significant difference for the well-differentiated group ( $HR = 2.74$ ,  $95\% \text{ CI} = 1.57\text{--}4.79$ ,  $P = 0.00022$ ; **Figure 4B**) and poor-differentiated group ( $HR = 4.1$ ,  $95\% \text{ CI} = 1.79\text{--}9.37$ ,  $P = 0.00035$ ; **Figure 4C**). **Figure 4D** indicates that patients with an MSS status in the high-risk group had importantly shorter

median OS than the low risk patients ( $HR = 3.83$ ,  $95\% \text{ CI} = 2.24\text{--}6.56$ ,  $P < 0.0001$ ). These results revealed that the prognostic ability of the 5-lncRNA signature was independent of the TNM stage, histologic grade and MSI status.

### Subgroup Analyses Based on the 5-lncRNA Signature According to Age and Gender

As we know, age is an important risk factor in the progress of carcinogenesis (35). We performed the subgroup analyses by separating these patients into  $\leq 60$ -year-old and  $\geq 60$ -year-old subgroups. Regardless of patients in the younger or older



groups, our 5-LncRNA signature still had the capacity to identify patients with different prognoses (Figures S1A,B). We also studied whether our 5-LncRNA signature was independent of gender by adopting similar methods as described above. In the entire TCGA PDAC cohort, the high-risk score of our signature significantly associated with an advert OS either in female or male patients ( $P = 0.0039$ ,  $P < 0.0001$ , respectively; Figures S1C,D). These results further demonstrates that our 5-LncRNA signature was definitely independent of age and gender.

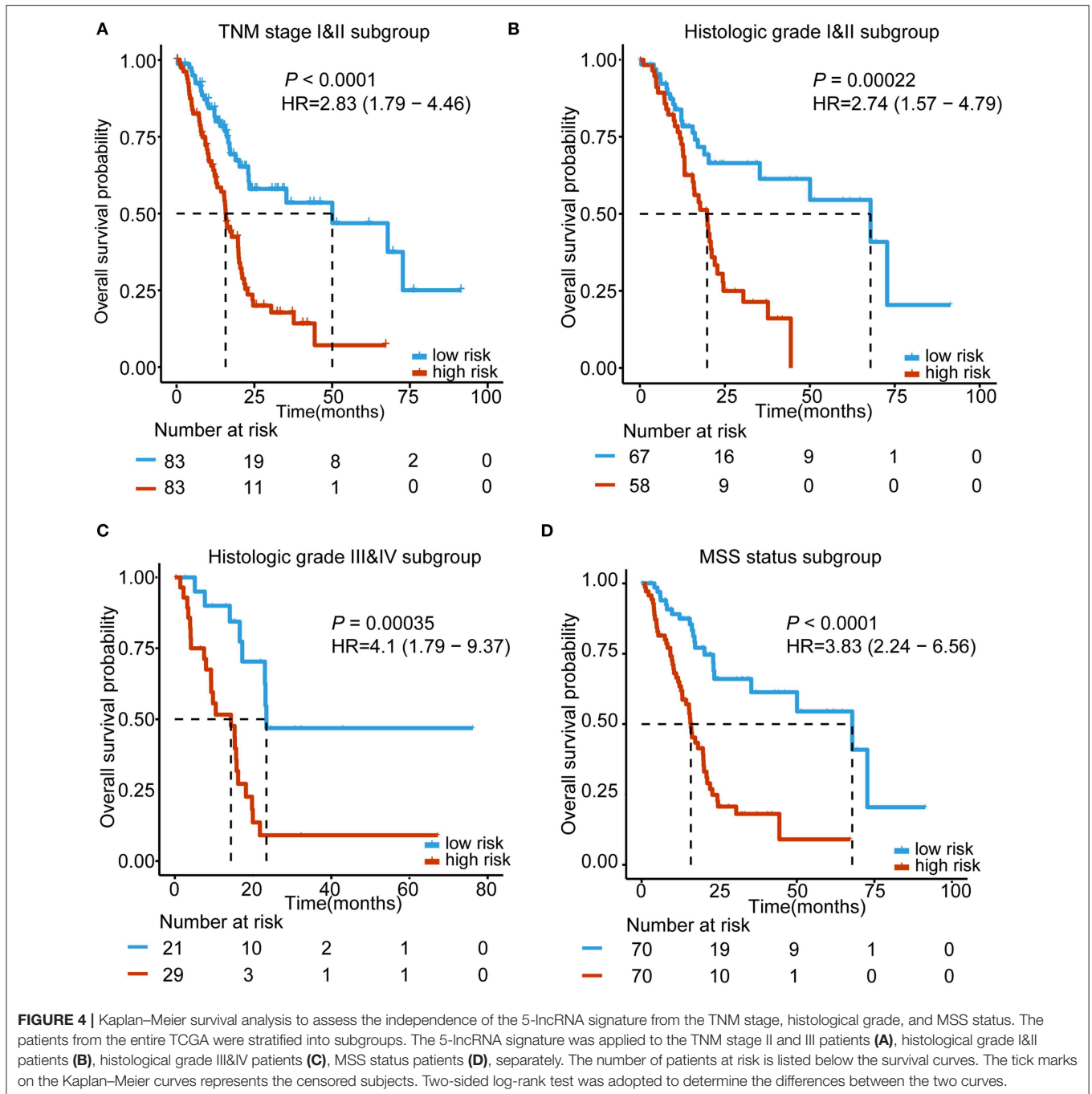
## ROC Analysis to Assess the Prognostic Value of the 5-LncRNA Signature

We conducted a ROC analysis to assess whether our 5-LncRNA signature could behave better than the present clinical parameters in predicting OS prognosis. As Figure 5A shows, the area under receiving operating curve (AUROC) of the 5-LncRNA risk score model was superior to those of TNM stage and histologic grade (0.68 vs. 0.53; 0.68 vs. 0.55). In addition, the AUROC of 5-LncRNA

risk score combined with TNM stage and histologic grade showed significant differences when compared to the signature alone (0.73 vs. 0.68). This phenomenon was also confirmed in our Fudan validation cohort (Figure 5B). Because there was no data related to the histologic grade of our patients, ROC analysis was performed according to the 5-LncRNA risk score and TNM stage. The results were similar to those in the entire TCGA cohort. The 5-LncRNA risk score possessed a better performance than TNM stage (0.70 vs. 0.62). What is more, when combined with our signature TNM stage, the combined model had a strong power for OS prediction (AUROC = 0.76).

## Enriched Functions and Pathways Associated With 5-LncRNA Signature

For the purpose of elucidating the 5-LncRNA related biological functions and pathways, we used cytoscape plug-in ClueGO and CluePedia to analyze the entire TCGA PDAC cohort according to the classification through our signature. Figure 6



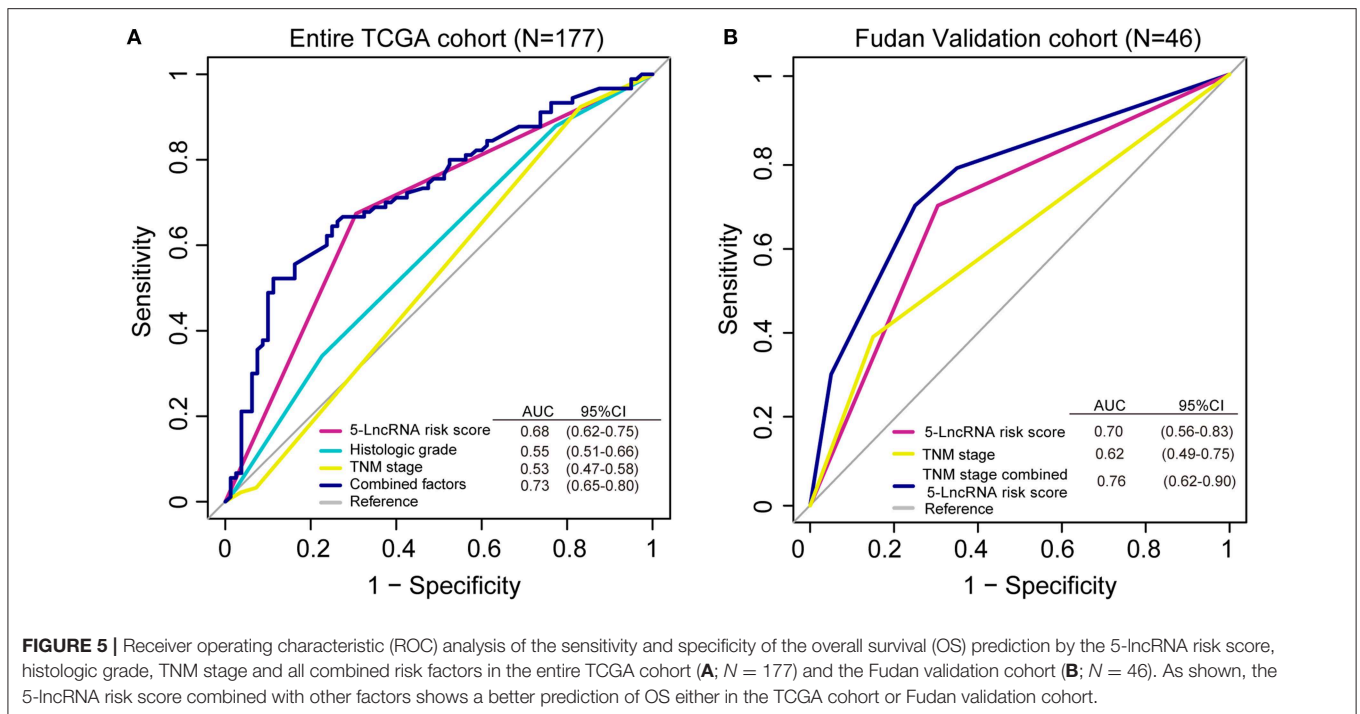
shows a set of enriched functions and pathways of the top 1,000 significantly DEGs in high vs. low risk PDAC patients in the TCGA dataset. In the analyses of the top 1,000 up-regulated DEGs in high risk groups, we found that a group of cancer-related pathways like programmed cell death, regulation of cell population proliferation, epithelial cell differentiation, cell adhesion, and the secretion of the pancreas were involved (Figure 6A). Similarly, several drug response pathways, signal transmission pathways including gated channel, negative regulation of synaptic transmission, and the cAMP signaling pathway, as well as the secretin receptors related to the

pancreas were also found to be changed in the pathways of the top 1,000 down-regulated DEGs in high risk groups (Figure 6B). Based on the above analyses, we propose that our 5-lncRNA signature could play a significant role in the carcinogenesis of PDAC.

## DISCUSSION

At present, with the emergence of many lncRNAs, researchers have altered their study attention from the traditional protein-coding genes to non-coding RNAs. Many publications nowadays





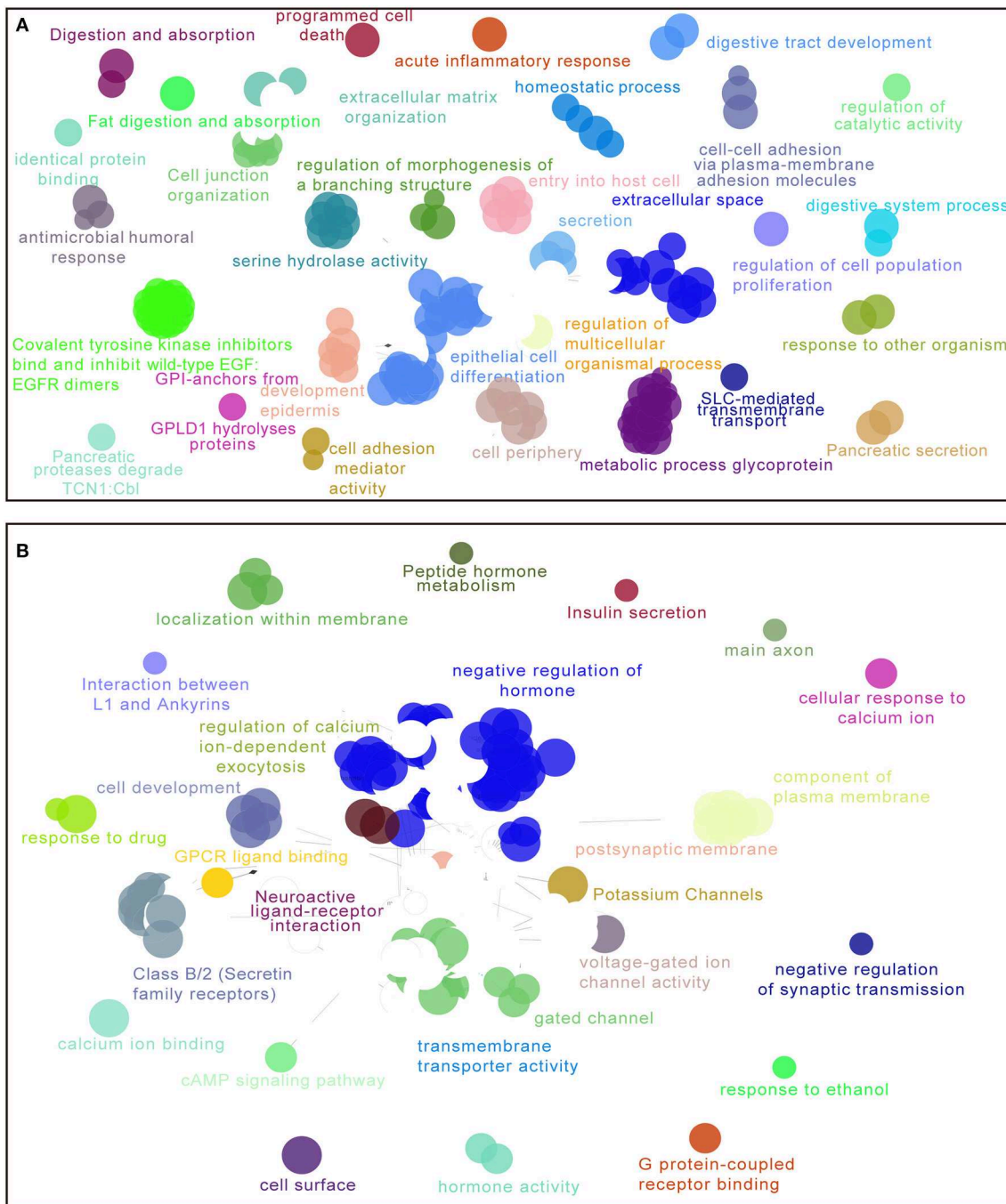
have revealed that lncRNA plays an important role in the process of tumor development. Zhang et al. and Song et al. have reported, respectively, that they found the multi-lncRNA signatures that can predict the prognosis of pancreatic cancer. However, several limitations exist in their studies. For example, their signatures have not been validated in an independent validation cohort (27, 28). In addition, Zhang et al. only used the Cox proportional hazards regression method to generate the model. This method was reported to be inappropriate for high-dimensional sequencing data when the ratio between the parameters and sample size is too low (36).

In our study, we constructed the PDAC lncRNA prognosis signature by mining the lncRNA data from the TCGA database using a LASSO algorithm. By investigating the relationship between the clinical data and lncRNA expression profiles of PDAC patients in TCGA training cohort, we established a novel 5-lncRNA signature that was significantly related to the OS. By utilizing this signature to the patients in the TCGA PDAC training cohort, statistical significance was seen in the survival curve between the high and low risk groups. The prognostic capacity of this 5-lncRNA signature could be internally verified in the TCGA PDAC validation cohort and our independent Fudan validation cohort, revealing the wide application and efficacy of this signature in predicting PDAC prognosis.

Moreover, we performed univariate and multivariate analyses to verify that our 5-lncRNA signature can be an independent risk variable in predicting PDAC OS. Although those clinical parameters like age, gender, histologic grade, and TNM stage did not have statistical significance in the TCGA cohort and Fudan validation cohort by multivariate analyses, we applied the subgroup analyses to further confirm the independence of our

5-lncRNA signature. As we know, it is widely agreed that TNM stage can work as an influential factor in predicting the prognosis of PDAC. TNM stage I&II generally indicates the early stage of PDAC, while TNM stage III&IV indicates the late stage. We classified all the TCGA PDAC patients into early stage and late stage stratum to promote further analysis. The stratified patients were successfully separated into high and low risk subgroups based on the 5-lncRNA signature, and there was an obvious split in the overall survival curve of TNM I&II between them. The number of TNM III&IV patients was little, so we did not further analyze this subgroup. Additionally, tumor histologic grade has also been reported to be an effective prognostic factor in PDAC (37, 38). Our 5-lncRNA signature could also successfully discriminate the high or low risk of patients in both the well and poor differentiated subgroup. MSI status was reported to be a significant predictor for resectable pancreatic cancer patients, and the prognosis of MSI-H status patients was better than MSS status patients (39). Given the terrible survival of MSS status patients and because the number of MSI-indeterminate (MSI-I)/MSI-low (MSI-L) patients was not enough for analysis, we tested the independence of our signature in the MSS status group. Patients of MSS status in the high-risk group were found to have shorter median OS than low risk patients.

In addition, there are many theories on aging. Some researchers think that the accumulation of genomic and epigenomic instability promotes cancer, while some think that tissue overstimulation and the hyper-activation of the DNA damage response causes cancer (40). PDAC was reported to be an age-dependent cancer (35). Our 5-lncRNA risk model maintained its powerful prognostic capacity when stratified by age distribution. Furthermore, there are reports that males have



**FIGURE 6** | Enriched functions and pathways of the top 1,000 significantly differentially expressed genes (DEGs) in high vs. low risk PDAC patients in the TCGA dataset. The interaction network was generated with the Cytoscape plug-in ClueGO and CluePedia. Functions and pathways of up-regulated DEGs (**A**), down-regulated DEGs (**B**). The size of the nodes shows the term significance after Bonferroni correction. The significant term of each group is highlighted.

higher susceptibility in cancer than females (41). By applying our risk model to the stratification analysis of gender, our 5-lncRNA signature also proved to be independent of gender.

To evaluate the predictive capacity of the 5-lncRNA signature, ROC analysis was carried out. In the diagnostic test, we can use AUROC to detect the accuracy and determine the predictive

value of biomarkers (42). ROC analysis demonstrated that this 5-lncRNA signature was superior to the TNM stage in PDAC OS evaluation (in the entire TCGA and Fudan validation cohort), and histologic grade (in the entire TCGA cohort). Interestingly, when we combined the 5-lncRNA risk score with TNM stage and histologic grade (if available), the prognostic ability was

better than any parameter alone in this model. The AUROC of the combined model reached 0.73 for the entire TCGA cohort, and 0.76 for the Fudan validation cohort, suggesting that it might enhance clinic-pathologic characteristics and strengthen the predictive accuracy of OS prognosis in PDAC.

Because the 5-lncRNA signature was able to separate high risk patients according to the risk score, we assumed that several biological functions and pathways related to this signature might affect the OS prognosis of PDAC. Nowadays, the main challenges for PDAC patients are metastatic problems and the lack of sensitive drugs for clinical treatment (1, 43), which seriously influence the prognosis of these patients. According to the analyses of cytoscape plug-in ClueGO and CluePedia, a number of cancer-related pathways were highlighted in the high-risk groups such as programmed cell death, regulation of cell population proliferation, epithelial cell differentiation, cell adhesion, and the secretion of pancreas pathway. Furthermore, several drug response pathways, signal transduction pathways, and the secretin receptors related to the pancreas were also found to be changed in the high-risk group. Therefore, these biological functions and pathways obtained from our 5-lncRNA signature might provide a strong backing for studying the molecular mechanisms and developing the potential targeted therapies.

The 5-lncRNA signature has been proven to be significantly connected with the overall survival of PDAC. However, the underlying function of these 5 lncRNAs have not been fully illuminated in this study. According to the negative coefficients and the downregulation of the five lncRNAs identified in high risk PDAC patients, we hypothesize that these lncRNAs are associated with better survival and may act as tumor suppressors in PDAC. RP11-159F24.5 was reported to be the antisense to nicotinamide nucleotide transhydrogenase (NNT) (44), and the deficiency of NNT will impede cell proliferation and tumorigenicity (45). RP11-744N12.2 is also named a smooth muscle and endothelial cell enriched migration/differentiation-associated long non-coding RNA (lncRNA SENCER) or FLII-AS1 (46). Lyu et al. demonstrated that SENCER interacted with CKAP4 to stabilize vascular endothelial cell adherens junctions (47). RP11-388M20.1 is the antisense to PYCARD, and PYCARD is reported to play an important role in promoting cellular apoptosis (48). In addition, the expression of RP11-356C4.5 is found to be up-regulated in colorectal tumor samples compared to normal control tissues (49). There are extremely rare reports regarding CTC-459F4.9. More studies about this lncRNA are needed in the near future.

Several limitations in our study should be pointed out. First, our 5-lncRNA signature was generated from the retrospective data, and more prospective datasets are needed to prove the clinical utility of our model. Second, owing to the limited number of patients recruited in our study, some subgroup analyses cannot be implemented. Third, there was no experimental data about the expression and mechanisms of the lncRNAs in PDAC samples, so more efforts should be invested to illuminate the potential association between our signature and the prognosis in PDAC.

## CONCLUSION

In summary, our study defined an innovative 5-lncRNA signature in PDAC. It is an integrated analysis of the available RNA-sequencing data and qPCR results from our own samples. The 5-lncRNA signature was proven to be independently associated with the OS of classical prognostic parameters and remains a good classifier in different subtypes of PDAC patients. This signature may provide insight into the prediction of PDAC prognosis. What is more, the obtained functions and pathways associated with our signature may facilitate the development of novel therapies for PDAC treatment.

## DATA AVAILABILITY STATEMENT

The level three RNA sequencing data and relevant clinical information of 177 PDAC patients were downloaded from TCGA database (<http://cancergenome.nih.gov/>). Other data generated or analyzed during this study are included in this published article and the **Supplementary Material**.

## ETHICS STATEMENT

Written informed consent was obtained from all our patients. The study was conducted in accordance with the Declaration of Helsinki, and the Ethical Committee of Huashan Hospital, Fudan University approved the study.

## AUTHOR CONTRIBUTIONS

HJ, NR, and QD: conceptualization and funding acquisition. CZ: data curation. JZ, XX, and YZhe: formal analysis. WC, MX, and YY: investigation. SW, QZ, and MA: methodology. FY, DF, LQ, and HJ: resources. NR and QD: supervision. CZ, SW, and QZ: writing—original draft. YZha, CB, HJ, NR, and QD: writing—review and editing. All authors read and approved the final manuscript.

## FUNDING

This work was financially supported by grants from the National Key Research and Development Program of China (2017YFC1308604), the National Natural Science Foundation of China (81472672, 81872356), the Shanghai International Science and Technology Collaboration Program (18410721900), and the China National Key Projects for Infectious Disease (2017ZX10203207004001).

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fonc.2019.01160/full#supplementary-material>

## REFERENCES

- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* (2019) 69:7–34. doi: 10.3322/caac.21551
- Rahib L, Smith BD, Aizenberg R, Rosenzweig AB, Fleshman JM, Matrisian LM. Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* (2014) 74:2913–21. doi: 10.1158/0008-5472.CAN-14-0155
- De Luca R, Blasi L, Alu M, Cristina V, Cicero G. Clinical efficacy of nab-paclitaxel in patients with metastatic pancreatic cancer. *Drug Des Dev Ther.* (2018) 12:1769–75. doi: 10.2147/DDDT.S165851
- Rombouts SJ, Vogel JA, van Santvoort HC, van Lienden KP, van Hillegersberg R, Busch OR, et al. Systematic review of innovative ablative therapies for the treatment of locally advanced pancreatic cancer. *Brit J Surg.* (2015) 102:182–93. doi: 10.1002/bjs.9716
- Garrido-Laguna I, Hidalgo M. Pancreatic cancer: from state-of-the-art treatments to promising novel therapies. *Nat Rev Clin Oncol.* (2015) 12:319–34. doi: 10.1038/nrclinonc.2015.53
- Lennon AM, Wolfgang CL, Canto MI, Klein AP, Herman JM, Goggins M, et al. The early detection of pancreatic cancer: what will it take to diagnose and treat curable pancreatic neoplasia? *Cancer Res.* (2014) 74:3381–9. doi: 10.1158/0008-5472.CAN-14-0734
- Kamisawa T, Wood LD, Itoi T, Takaori K. Pancreatic cancer. *Lancet.* (2016) 388:73–85. doi: 10.1016/S0140-6736(16)00141-0
- Heestand GM, Murphy JD, Lowy AM. Approach to patients with pancreatic cancer without detectable metastases. *J Clin Oncol.* (2015) 33:1770–8. doi: 10.1200/JCO.2014.59.7930
- Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet.* (2009) 10:155–9. doi: 10.1038/nrg2521
- Lipovich L, Johnson R, Lin CY. MacroRNA underdogs in a microRNA world: evolutionary, regulatory, and biomedical significance of mammalian long non-protein-coding RNA. *Biochim Biophys Acta.* (2010) 1799:597–615. doi: 10.1016/j.bbtagm.2010.10.001
- Consortium EP, Birney E, Stamatoyannopoulos JA, Dutta A, Guigo R, Gingeras TR, et al. Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature.* (2007) 447:799–816. doi: 10.1038/nature05874
- Khanduja JS, Calvo IA, Joh RL, Hill IT, Motamedi M. Nuclear Noncoding RNAs and Genome Stability. *Mol Cell.* (2016) 63:7–20. doi: 10.1016/j.molcel.2016.06.011
- Kornienko AE, Guenzl PM, Barlow DP, Pauler FM. Gene regulation by the act of long non-coding RNA transcription. *BMC Biol.* (2013) 11:59. doi: 10.1186/1741-7007-11-59
- Fatica A, Bozzoni I. Long non-coding RNAs: new players in cell differentiation and development. *Nat Rev Genet.* (2014) 15:7–21. doi: 10.1038/nrg3606
- Ulitsky I, Bartel DP. lincRNAs: genomics, evolution, and mechanisms. *Cell.* (2013) 154:26–46. doi: 10.1016/j.cell.2013.06.020
- Bolha L, Ravnik-Glavac M, Glavac D. Long noncoding RNAs as biomarkers in cancer. *Dis Markers.* (2017) 2017:7243968. doi: 10.1155/2017/7243968
- Peng JF, Zhuang YY, Huang FT, Zhang SN. Noncoding RNAs and pancreatic cancer. *World J Gastroenterol.* (2016) 22:801–14. doi: 10.3748/wjg.v22.i2.801
- Shi X, Guo X, Li X, Wang M, Qin R. Loss of Linc01060 induces pancreatic cancer progression through vinculin-mediated focal adhesion turnover. *Cancer Lett.* (2018) 433:76–85. doi: 10.1016/j.canlet.2018.06.015
- Hui B, Ji H, Xu Y, Wang J, Ma Z, Zhang C, et al. RREB1-induced upregulation of the lncRNA AGAP2-AS1 regulates the proliferation and migration of pancreatic cancer partly through suppressing ANKRD1 and ANGPTL4. *Cell Death Dis.* (2019) 10:207. doi: 10.1038/s41419-019-1384-9
- Chen S, Chen JZ, Zhang JQ, Chen HX, Qiu FN, Yan ML, et al. Silencing of long noncoding RNA LINC00958 prevents tumor initiation of pancreatic cancer by acting as a sponge of microRNA-330-5p to down-regulate PAX8. *Cancer Lett.* (2019) 446:49–61. doi: 10.1016/j.canlet.2018.12.017
- Stratford JK, Bentrem DJ, Anderson JM, Fan C, Volmar KA, Marron JS, et al. A six-gene signature predicts survival of patients with localized pancreatic ductal adenocarcinoma. *PLoS Med.* (2010) 7:e1000307. doi: 10.1371/journal.pmed.1000307
- Serghiou S, Kyriakopoulou A, Ioannidis JP. Long noncoding RNAs as novel predictors of survival in human cancer: a systematic review and meta-analysis. *Mol Cancer.* (2016) 15:50. doi: 10.1186/s12943-016-0535-1
- Zhou M, Zhao H, Xu W, Bao S, Cheng L, Sun J. Discovery and validation of immune-associated long non-coding RNA biomarkers associated with clinically molecular subtype and prognosis in diffuse large B cell lymphoma. *Mol Cancer.* (2017) 16:16. doi: 10.1186/s12943-017-0580-4
- Jin K, Luo G, Xiao Z, Liu Z, Liu C, Ji S, et al. Noncoding RNAs as potential biomarkers to predict the outcome in pancreatic cancer. *Drug Des Dev Ther.* (2015) 9:1247–55. doi: 10.2147/DDDT.S77597
- van 't Veer LJ, Dai H, van de Vijver MJ, He YD, Hart AA, Mao M, et al. Gene expression profiling predicts clinical outcome of breast cancer. *Nature.* (2002) 415:530–6. doi: 10.1038/415530a
- Birnbaum DJ, Finetti P, Lopresti A, Gilabert M, Poizat F, Raoul JL, et al. A 25-gene classifier predicts overall survival in resectable pancreatic cancer. *BMC Med.* (2017) 15:170. doi: 10.1186/s12916-017-0936-z
- Zhang H, Zhu M, Du Y, Zhang H, Zhang Q, Liu Q, et al. A panel of 12-lncRNA signature predicts survival of pancreatic adenocarcinoma. *J Cancer.* (2019) 10:1550–9. doi: 10.7150/jca.27823
- Song J, Xu Q, Zhang H, Yin X, Zhu C, Zhao K, et al. Five key lncRNAs considered as prognostic targets for predicting pancreatic ductal adenocarcinoma. *J Cell Biochem.* (2018) 119:4559–69. doi: 10.1002/jcb.26598
- Tibshirani R. The lasso method for variable selection in the Cox model. *Stat Med.* (1997) 16:385–95.
- Goeman JJ. L1 penalized estimation in the Cox proportional hazards model. *Biomet J Biomet Z.* (2010) 52:70–84. doi: 10.1002/bimj.200900028
- Gentleman RC, Carey VJ, Bates DM, Bolstad B, Dettling M, Dudoit S, et al. Bioconductor: open software development for computational biology and bioinformatics. *Genome Biol.* (2004) 5:R80. doi: 10.1186/gb-2004-5-10-r80
- Jiao X, Sherman BT, Huang da W, Stephens R, Baseler MW, Lane HC, et al. DAVID-WS: a stateful web service to facilitate gene/protein list analysis. *Bioinformatics.* (2012) 28:1805–6. doi: 10.1093/bioinformatics/bts251
- Bindea G, Mlecnik B, Hackl H, Charoentong P, Tosolini M, Kirilovsky A, et al. ClueGO: a Cytoscape plug-in to decipher functionally grouped gene ontology and pathway annotation networks. *Bioinformatics.* (2009) 25:1091–3. doi: 10.1093/bioinformatics/btp101
- Mlecnik B, Bindea G, Angell HK, Maby P, Angelova M, Tougeron D, et al. Integrative analyses of colorectal cancer show immunoscore is a stronger predictor of patient survival than microsatellite instability. *Immunity.* (2016) 44:698–711. doi: 10.1016/j.immuni.2016.02.025
- Liszka L, Pajak J, Mrowiec S, Zielinska-Pajak E, Lampe P, Golka D. Age distribution patterns of patients with conventional ductal adenocarcinoma of the pancreas. A single-institution study of 580 cases re-evaluated using current histopathological diagnostic criteria. *Polish J Pathol.* (2010) 61:65–77.
- Simon R, Altman DG. Statistical aspects of prognostic factor studies in oncology. *Brit J Cancer.* (1994) 69:979–85. doi: 10.1038/bjc.1994.192
- Luttges J, Schemm S, Vogel I, Hedderich J, Kremer B, Kloppel G. The grade of pancreatic ductal carcinoma is an independent prognostic factor and is superior to the immunohistochemical assessment of proliferation. *J Pathol.* (2000) 191:154–61. doi: 10.1002/(SICI)1096-9896(200006)191:2<154::AID-PATH603>3.0.CO;2-C
- Macias N, Sayagues JM, Esteban C, Iglesias M, Gonzalez LM, Quinones-Sampedro J, et al. Histologic tumor grade and preoperative biliary drainage are the unique independent prognostic factors of survival in pancreatic ductal adenocarcinoma patients after pancreaticoduodenectomy. *J Clin Gastroenterol.* (2018) 52:e11–7. doi: 10.1097/MCG.0000000000000793
- Nakata B, Wang YQ, Yashiro M, Nishioka N, Tanaka H, Ohira M, et al. Prognostic value of microsatellite instability in resectable pancreatic cancer. *Clin Cancer Res.* (2002) 8:2536–40.
- Falandry C, Bonnefoy M, Freyer G, Gilson E. Biology of cancer and aging: a complex association with cellular senescence. *J Clin Oncol.* (2014) 32:2604–10. doi: 10.1200/JCO.2014.55.1432
- Dorak MT, Karpuzoglu E. Gender differences in cancer susceptibility: an inadequately addressed issue. *Front Genet.* (2012) 3:268. doi: 10.3389/fgene.2012.00268

42. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*. (1982) 143:29–36. doi: 10.1148/radiology.143.1.7063747
43. Taieb J, Pointet AL, Van Laethem JL, Laquente B, Pernot S, Lordick F, et al. What treatment in 2017 for inoperable pancreatic cancers? *Ann Oncol*. (2017) 28:1473–83. doi: 10.1093/annonc/mdx174
44. Xu Z, Zhou X, Li H, Chen Q, Chen G. Identification of the key genes and long noncoding RNAs in ankylosing spondylitis using RNA sequencing. *Int J Mol Med*. (2019) 43:1179–92. doi: 10.3892/ijmm.2018.4038
45. Ho HY, Lin YT, Lin G, Wu PR, Cheng ML. Nicotinamide nucleotide transhydrogenase (NNT) deficiency dysregulates mitochondrial retrograde signaling and impedes proliferation. *Redox Biol*. (2017) 12:916–28. doi: 10.1016/j.redox.2017.04.035
46. Boulberdaa M, Scott E, Ballantyne M, Garcia R, Descamps B, Angelini GD, et al. A role for the long noncoding RNA SENCR in commitment and function of endothelial cells. *Mol Ther*. (2016) 24:978–90. doi: 10.1038/mt.2016.41
47. Lyu Q, Xu S, Lyu Y, Choi M, Christie CK, Slivano OJ, et al. SENCR stabilizes vascular endothelial cell adherens junctions through interaction with CKAP4. *Proc Natl Acad Sci USA*. (2019) 116:546–55. doi: 10.1073/pnas.1810729116
48. Deswaerte V, Nguyen P, West A, Browning AF, Yu L, Ruwanpura SM, et al. Inflammasome adaptor ASC suppresses apoptosis of gastric cancer cells by an IL18-mediated inflammation-independent mechanism. *Cancer Res*. (2018) 78:1293–307. doi: 10.1158/0008-5472.CAN-17-1887
49. Xiao WH, Qu XL, Li XM, Sun YL, Zhao HX, Wang S, et al. Identification of commonly dysregulated genes in colorectal cancer by integrating analysis of RNA-Seq data and qRT-PCR validation. *Cancer Gene Ther*. (2015) 22:278–84. doi: 10.1038/cgt.2015.20

**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.

Copyright © 2019 Zhou, Wang, Zhou, Zhao, Xia, Chen, Zheng, Xue, Yang, Fu, Yin, Atyah, Qin, Zhao, Bruns, Jia, Ren and Dong. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). 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.