

A five-long non-coding RNA signature with the ability to predict overall survival of patients with lung adenocarcinoma

LIZHONG ZENG, WEI WANG, YANG CHEN, XIN LV,
JINGYAN YUAN, RUIYING SUN and SHUANYING YANG

Department of Respiratory, The Second Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shanxi 710000, P.R. China

Received April 23, 2019; Accepted September 24, 2019

DOI: 10.3892/etm.2019.8138

Abstract. An increasing number of studies have indicated that the abnormal expression of certain long non-coding RNAs (lncRNAs) is linked to the overall survival (OS) of patients with lung adenocarcinoma (LUAD). The aim of the present study was to establish an lncRNA signature to predict the survival of patients with LUAD. The gene expression profiles and associated clinical information of patients with LUAD were downloaded from The Cancer Genome Atlas database. The cohort was randomly sub-divided into training and verification cohorts. Univariate Cox regression analysis was performed on differentially expressed lncRNAs in the training cohort to select candidate lncRNAs closely associated with survival. Next, a risk score (RS) model consisting of 5 lncRNAs was established by multivariate Cox regression analysis on candidate lncRNAs. Using the median RS obtained from the training cohort as a cut-off point, patients were classified into high- and low-risk groups. Kaplan-Meier survival analysis revealed a significant difference in OS between high- and low-risk groups. The survival prediction ability of the 5-lncRNA signature was further tested in the verification and total cohorts. The results proved that the 5-lncRNA signature had good reliability and stability in survival prediction for patients with LUAD. The univariate Cox regression analysis for the 5-lncRNA signature in each cohort indicated that the 5-lncRNA signature was

closely associated with survival. Multivariate Cox regression analysis and stratification analysis proved that the prognostic signature was an independent predictor of survival for patients with LUAD. In addition, functional enrichment analysis indicated that the 5 prognostic lncRNAs may be involved in the tumorigenesis of LUAD through cancer-associated pathways and biological processes. Taken together, the present study provided a 5-lncRNA signature that may serve as an independent survival predictor for patients with LUAD.

Introduction

Lung cancer, the biggest contributor to cancer-associated mortality worldwide, has a severe impact on public health (1). At present, the 5-year survival rate for lung cancer is only 18%, which is in contrast with the stable increase in the survival rates of other types of cancer (2). Among all lung cancer subtypes, non-small cell lung cancer (NSCLC) accounts for ~85%. Of note, lung adenocarcinoma (LUAD) is the most common and malignant pathological type of NSCLC (3). Although the 5-year survival rate of LUAD has increased with the emergence of targeted drugs, the mortality rate remains high due to the lack of effective diagnostic and prognostic biomarkers (4,5). To reduce the mortality rate of patients with LUAD, it is therefore essential to establish a good prognostic signature to guide the patients' treatment and clinical management.

It has been reported that ~90% of mammalian genomes are transcribed into non-coding RNA (ncRNA) (6). Long ncRNAs (lncRNAs) are one of the most important classes of the ncRNA family and are >200 nucleotides in length (7). Accumulating studies have suggested that certain lncRNAs have mutations or changes in expression levels in various types of cancer (8-10). Differentially expressed lncRNAs are involved in the aberrant processes of cancer, including cell cycle, apoptosis and chemoresistance (11). These cancer-associated lncRNAs have a critical role in tumorigenesis and metastasis through different mechanisms, including regulating the gene expression by serving as a guide to target chromatin-modifying complexes to a specific gene location or acting as a competing endogenous (ce)RNA that competitively binds to microRNA (miR) to regulate gene expression (12). With the importance of lncRNAs in cancer recognized, an increasing number of studies have explored the function of lncRNA in NSCLC. For instance, certain significant lncRNAs with a gene regulatory mechanism in LUAD

Correspondence to: Professor Shuanying Yang, Department of Respiratory, The Second Affiliated Hospital of Xi'an Jiaotong University, 157 Xiwu Road, Xincheng, Xi'an, Shanxi 710000, P.R. China
E-mail: yangshuanying66@163.com

Abbreviations: OS, overall survival; LUAD, lung adenocarcinoma; TCGA, the cancer genome atlas; RS, risk score; lncRNA, long non-coding RNA; NSCLC, non-small-cell lung cancer; GEO, Gene Expression Omnibus; AIC, akaike information criterion; ROC, receiver operating characteristic; AUC, area under the ROC curve; HR, hazard ratio; RT-qPCR, reverse transcription-quantitative PCR; PCGs, protein-coding genes

Key words: long non-coding RNA, overall survival, lung adenocarcinoma, signature

were discovered through pathway crosstalk analysis (13) and several of them were identified to have the potential to act as therapeutic targets or diagnostic markers for NSCLC (14-16). It is also critical to explore the association between lncRNAs and cancer prognosis. To date, lncRNA-associated prognostic signatures have been established in several types of cancer, including gastric cancer (17), urothelial carcinoma (18) and hepatocellular carcinoma (19). Numerous prognostic lncRNA signatures have been reported for various lung cancer subtypes, including lung squamous cell carcinoma (20), NSCLC (21) and NSCLC in elderly subjects (22). Of course, numerous studies have explored the association between lncRNA and the prognosis of LUAD. lncRNAs serving as prognostic biomarkers for LUAD were identified through a ceRNA network analysis (23). A prognostic signature based on lncRNA suggested the existence of a tumor protein 53-dependent subtype of LUAD with poor survival (24). Prognostic signatures for LUAD constructed from lncRNA or lncRNA combined with mRNA have also been reported (25,26). However, to date, no ideal lncRNA signature has been developed for use in the clinical setting, possibly due to limited sample size and lack of systematic investigation. The lncRNA signature established in the present study may provide a reference value for building the prognostic signature that may be applied in the clinic.

In the present study, a 5-lncRNA signature with good reliability and stability in the prognostication of patients with LUAD was successfully established. The predictive ability of the prognostic signature was independent of other clinical factors. The results of the functional enrichment analysis demonstrated that the 5 lncRNAs may be involved in the tumorigenesis of LUAD. In short, the present study provided an lncRNA signature that may be utilized for survival prediction of patients with LUAD.

Materials and methods

LUAD database and clinical information of patients. The lncRNA expression profiles of 535 LUAD tumor and 49 non-tumor tissues, as well as the clinical information of 504 patients with LUAD, were downloaded from the official website of The Cancer Genome Atlas (TCGA; <https://gdc-portal.nci.nih.gov/>). After excluding those patients with incomplete clinical information, the data of 486 patients were retained for analysis in the present study. The correlative clinical information included overall survival (OS), age, sex and TNM stage. The 486 patients were randomly divided into two cohorts. Of these, 264 LUAD patients were used as a training cohort to build a prognostic signature, while the 222 remaining LUAD patients were used as the verifying cohort to test the prognostic ability of the signature. Detailed information on the cohorts used in the present study is provided in Table I.

Acquisition and processing of lncRNA expression profiles for LUAD patients. The expression profiles were acquired for 17,109 mRNAs and 1,787 lncRNAs after adding annotation using the Ensemble database (<http://asia.ensembl.org/index.html>). Next, the R package 'edgeR' (27) was used to log₂-transform the RNA-sequencing expression values and normalize the data and differential expression analysis, using log₂|fold change|>1 and adjusted P<0.05 as the threshold

to screen out differentially expressed RNAs. A total of 841 differentially expressed lncRNAs were screened.

Acquisition and processing of lncRNA expression profiles for LUAD from the gene expression omnibus (GEO) database. A total of three GEO datasets were downloaded, which contained gene expression, clinical information and the platform annotation file (GPL570-Affymetrix Human Genome U133 Plus 2.0 Array) from the National Center for Biotechnology Information (NCBI) GEO database (<http://www.ncbi.nlm.nih.gov/geo>). The average expression value of a gene was used when it corresponded to multiple probes. After converting the probe information to gene symbols, the three datasets were merged and the expression levels of different batches were normalized using the R package 'sva'. A total of 418 patients with LUAD were included in the present study, comprising 204 patients from the dataset GSE31210, 85 patients from GSE30219 and 129 patients from GSE50081.

Construction of a prognostic lncRNA signature. In the training cohort, univariate Cox regression analysis was used to evaluate the association between the expression of each differentially expressed lncRNA and the OS of patients with LUAD. Considering the number of lncRNAs selected and their association with prognosis, those lncRNAs with P<0.001 were considered as candidate lncRNAs. These candidate lncRNAs were further subjected to multivariate Cox regression analysis to select a set of lncRNAs, thereby establishing an RS model. The multivariate Cox regression analysis was performed using a mathematical model based on the Akaike Information Criterion (AIC) (28). The model based on the AIC was used to construct a prognostic signature with the best predictive ability but the least number of lncRNAs. The calculation formula of the RS was as follows:

$$RS = \sum_{i=1}^N (Exp_i \times Coe_i)$$

where N represents the total number of prognostic lncRNAs, *Exp_i* the expression of a certain lncRNA and *Coe_i* the regression coefficient obtained from the multivariate Cox regression analysis for a certain lncRNA numbered as *i*. Based on this equation, each patient with LUAD had an RS and the median RS was treated as a cut-off point to stratify the patients into low- and high-risk groups. Univariate and multivariate Cox regression analysis was performed using the Survival R package from the CRAN package repository (<https://cran.r-project.org/web/packages/>).

Statistical analysis. Kaplan-Meier survival analysis and two-sided log-rank tests were used to evaluate the difference in OS between low- and high-risk groups in each cohort. The Kaplan-Meier survival analysis was performed using the Survival R package. After the patients with LUAD were divided into multiple groups according to a certain clinical factor, the Kruskal-Wallis test was used to analyze whether a certain lncRNA was significantly influenced by this clinical factor. The accuracy of the prognostic signature in predicting the 5-year survival rate was assessed using time-dependent receiver operating characteristic (ROC) analysis. ROC analysis was performed using the R package 'survivalROC' (29). The area under the ROC curve (AUC) was calculated. Univariate Cox regression analysis was performed in each cohort to test whether the prognostic signature was associated with the

Table I. Clinical characteristics of lung adenocarcinoma patients in the present study.

Characteristics	Training cohort (n=264)	Verifying cohort (n=222)	Total cohort (n=486)
Vital status			
Alive	163 (61.7)	145 (65.3)	308 (63.4)
Dead	101 (38.3)	77 (34.7)	178 (36.6)
Age (years)			
<65	121 (45.8)	95 (42.8)	216 (44.4)
≥65	143 (54.2)	127 (57.2)	270 (55.6)
Sex			
Female	143 (54.2)	120 (54.1)	263 (54.1)
Male	121 (45.8)	102 (45.9)	223 (45.9)
Stage			
I	143 (54.2)	120 (54.1)	263 (54.1)
II	58 (22.0)	59 (26.6)	117 (24.1)
III/IV	63 (23.8)	43 (19.3)	106 (21.8)

Values are expressed as n (%).

survival of patients with LUAD. Multivariate Cox regression and stratification analysis were further performed in each cohort to test whether the prognostic signature had an independent predictive value regarding survival. At the same time, the hazard ratio (HR) and 95% confidence interval (CI) were determined. At RT-qPCR data were analyzed using GraphPad Prism 8.0 (GraphPad Software, Inc.). Significant differences of reverse transcription-quantitative PCR (RT-qPCR) data were analyzed using an unpaired Student's t-test, and $P < 0.05$ was considered to indicate a statistically significant difference.

Predicting the functions and pathways of lncRNAs. The Pearson's correlation coefficients for the correlation between the prognostic lncRNAs and protein-coding genes were calculated in order to select co-expressed PCGs. Genes significantly co-expressed with at least one of the prognostic lncRNAs were treated as lncRNA-associated PCGs (|Pearson's correlation coefficient| > 0.40 and $P < 0.01$). Gene ontology (GO) enrichment analysis for the co-expressed PCGs was performed using the Database for Annotation, Visualization and Integrated Discovery (DAVID) bioinformatics tool (<https://david.ncifcrf.gov/>) (30) and the GO terms were limited to 'Biological Process'. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis for the co-expressed PCGs was performed using the KO-based Annotation System bioinformatics tool (<http://kobas.cbi.pku.edu.cn/>) (31) with the entire human genome as the background. The results of the functional enrichment analysis with a $P < 0.005$ were regarded as the potential biological functions. Enriched GO terms with similar function were clustered and the major categories of clustering were visualized through the Enrichment Map plugin (32) in Cytoscape (33). Significantly enriched KEGG pathways were visualized with the R package 'ggplot2' (34).

RT-qPCR verification of lncRNA expression in LUAD tissues and cell lines. A total of 14 paired LUAD and adjacent non-tumor tissue samples were used in the present study,

which were collected at the Second Affiliated Hospital of Xian Jiaotong University (Xian, China) between March 2017 and December 2018. The study protocol was approved by the Ethics Committee of the Second Affiliated Hospital of Xian Jiaotong University (Xian, China) and ethics consent had been received written informed consents from all participating patients. Detailed clinical information of these patients are presented in Table SI. The human lung cancer cell lines A549, PC-9 and H1299 and the normal human lung epithelial cell line BEAS-2B were purchased from the Type Culture Collection Cell Bank of the Chinese Academy of Sciences (Shanghai, China). All cell lines were cultured in RPMI-1640 (Hyclone; GE Healthcare Life Sciences) containing 10% FBS at 37°C in a 5% CO₂ incubator. Total RNA was extracted from lung tissues using Fast1000 (Xfyangbio) and from cell lines using Fast 200 (Xfyangbio). The total RNA was reverse-transcribed using the PrimerScript™ RT reagent kit (Takara Biotechnology Co., Ltd.). The reverse transcription mixture was incubated at 37°C for 15 min and 85°C for 5 sec. The TB Green® Premix Ex Taq™ II (Takara Biotechnology Co., Ltd.) was used for detecting the gene amplification and qPCR was performed on the CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad Laboratories, Inc.). The qPCR reaction mixture was incubated at 95°C for 30 sec, followed by 40 cycles of 95°C for 3 sec and 60°C for 30 sec. The PCR primers for the 5 lncRNAs are listed in Supplemental Table SII. All experimental procedures were performed according to the manufacturer's protocol and all reactions were performed in triplicate. The 2^{-ΔΔC_q} method (35) was used to calculate the expression of lncRNAs and β-actin was used as an internal reference.

Results

Construction of a 5-lncRNA signature to predict the OS of patients with LUAD. First, each of the 841 differentially expressed lncRNAs was subjected to univariate Cox regression analysis in the training cohort. A total of 7 lncRNAs were

selected as candidate lncRNAs ($P < 0.001$). These candidate lncRNAs were then further subjected to multivariate Cox regression analysis. Finally, five lncRNAs, including neuropeptide S receptor 1-antisense RNA 1 (NPSR1-AS1), opioid growth factor receptor pseudogene 1 (OGFRP1), integrin subunit beta 1 divergent transcript (ITGB1-DT), LIM domain 7 downstream neighbor (LMOTDN) and protein kinase cyclic GMP-dependent 1-antisense RNA 1 (PRKG1-AS1), were selected to construct a risk model. According to the results of the multivariate Cox regression analysis, an RS model was successfully established, as per the following equation: $RS = (0.155 \times \text{NPSR1-AS1 expression}) + (0.419 \times \text{OGFRP1 expression}) + (0.109 \times \text{ITGB1-DT expression}) + (-0.186 \times \text{LMOTDN expression}) + (0.151 \times \text{PRKG1-AS1 expression})$. The 5 lncRNAs were differentially expressed in 535 tumor vs. 49 non-tumor tissues ($P < 0.0005$; Fig. S1). Detailed information about the 5 lncRNAs is provided in Table II.

The 5-lncRNA signature predicted the OS of patients with LUAD in the training cohort. Based on the RS equation, the RS of each patient with LUAD was calculated in the training cohort. The results suggested that the median RS was 2.593. The 264 patients with LUAD in the training cohort were divided into high- ($n=132$) and low-risk groups ($n=132$) with $RS=2.593$ as the cut-off point. The Kaplan-Meier survival analysis indicated that the median survival of LUAD patients in the high-risk group was significantly shorter (1.63 years) than that in the low-risk group (1.91 years; $P=1.418 \times 10^{-6}$; Fig. 1A). Specifically, the 3-, 5- and 8-year survival rates were 47.5, 6.2 and 0.0% in the high-risk group, and 68.9, 45.1 and 31.6% in the low-risk group, respectively. From the time-dependent ROC curve, it was determined that the AUC value was 0.784 for the 5-lncRNA signature to predict the 5-year survival rate in the training cohort, indicating excellent reliability of the prognostic signature in predicting survival. The ROC curve is presented in Fig. 1B. The risk distribution and vital status of 264 LUAD patients from the training cohort are presented in Fig. 1C, and heatmap of the 5 lncRNAs expression profiles in training cohort is presented in Fig. 1D. Of these 5 lncRNAs, high expression of NPAR1-AS1, OGFRP1, ITGB1-DT and PRKG1-AS1 indicated to be associated with high RSs, while high expression of LMOTDN was associated with a low RS. Furthermore, the number of mortalities in the low-risk group was lower than that in the high-risk group. Univariate Cox regression analysis was performed on the 5-lncRNA signature in the training cohort. The results suggested that the prognostic signature was significantly linked to the survival of LUAD patients ($P < 0.001$, $HR=2.743$, 95% $CI=1.792-4.200$). More detailed results are provided in Table III.

Evaluation of the advantages of the 5-lncRNA signature. To evaluate the possible advantages of the 5-lncRNA signature in predicting the survival of LUAD patients, the same data and methods as those above were used to analyze the differentially expressed mRNAs. A total of 25 mRNAs significantly associated with OS ($P < 0.001$) were obtained. The top 7 mRNAs were subjected to multivariate Cox analysis and a 5-mRNA signature, including family with sequence similarity 189 member A2, collagen type XXII alpha 1 chain (COL22A1),

Clq and tumor necrosis factor related 6 (C1QTNF6), neurotensin receptor 1 (NTSR1) and cell death inducing DNA fragmentation factor subunit alpha like effector c, was obtained. The AUC value of the 5-mRNA signature to predict the 5-year survival rate of LUAD patients was 0.726 (Fig. S2). Subsequently, the same data and methods were used to analyze the differentially expressed lncRNA and mRNA together, and a total of 32 genes closely linked to survival were obtained ($P < 0.001$). The top 7 genes were subjected to multivariate COX analysis, and a 4-gene signature consisting of 2 lncRNAs (OGFRP1 and LINC01322) and 2 mRNAs (COL22A1 and NTSR1) was obtained. The AUC value of the 4-gene signature in predicting the 5-year survival of LUAD patients was 0.738 (Fig. S3). It is worth mentioning that 16 of the 25 mRNAs significantly associated with the prognosis were included in the co-expressed PCGs of the 5 lncRNAs obtained in the present study. The 5-lncRNA signature had a higher AUC value compared with the 5-mRNA and 4-gene signatures, which indicated that 5-lncRNA signature was more reliable in predicting prognosis. Considering the accuracy and complexity of the predictive model, the 5-lncRNA signature had certain advantages in predicting survival.

Verification of the ability of the 5-lncRNA signature to predict the OS of patients with LUAD. The ability of the 5-lncRNA signature to predict survival was further assessed in the verification and total cohorts. Based on the RS equation, the RS of each patient with LUAD in the verification and total cohort was calculated. With the median RS of 2.539 as the cut-off point, 222 LUAD patients in the verification cohort and 486 LUAD patients in the total cohort were classified into high- and low-risk groups ($n=113$ and 109, and $n=245$ and 241, respectively). The results of the Kaplan-Meier survival analysis for the verification and total cohorts were consistent with those in for training cohort. In the verification cohort, the median survival of LUAD patients in the high- and low-risk group was 1.66 and 2.12 years, respectively ($P=3.861 \times 10^{-3}$; Fig. 2A). The 3-, 5- and 8-year survival rates were 44.3, 25.9 and 0.0% in the high-risk group, and 71.0, 45.1 and 30.1% in the low-risk group, respectively. In the total cohort, the median survival of patients with LUAD in the high- and low-risk groups was 1.64 and 1.95 years, respectively ($P=1.724 \times 10^{-7}$; Fig. 2B). The 3-, 5- and 8-year survival rates were 47.3, 22.4 and 0.0% in the high-risk group, and 71.0, 47.4 and 37.6% in the low-risk group, respectively.

The risk distribution and vital status in LUAD patients from the verification and total cohorts are presented in Fig. 2C and D, the heatmap of the 5 lncRNAs expression profiles in verification and total cohorts are presented in Fig. 2E and F, respectively. As expected, the patients with high-risk LUAD had the tendency to express the high-risk lncRNAs, while the protective lncRNA was upregulated in low-risk patients. Mortality rate in the high-risk group was higher than that in the low-risk group (verification cohort: 41.6% vs. 27.5%; total cohort: 44.9% vs. 28.2%). Univariate Cox regression analysis was performed on the 5-lncRNA signature in these two cohorts. Similar results to those obtained in the training cohort were obtained: The prognostic signature was closely associated with survival (verification cohort: $P=0.005$, $HR=1.962$, 95% $CI=1.232-3.127$; total cohort: $P < 0.001$, $HR=2.225$, 95%

Table II. The five long non-coding RNAs significantly associated with prognosis of lung adenocarcinoma patients in the training cohort.

Gene name	Ensemble ID	Chromosomal position	Hazard ratio ^a	Coefficient ^a	P-value ^a
OGFRP1	ENSG00000182057.4	chr22: 42,269,753-42,275,196	1.52 (1.18-1.97)	0.419	0.001
ITGB1-DT	ENSG00000229656	chr10: 32,958,845-33,082,102	1.12 (1.01-1.23)	0.109	0.023
LMO7DN	ENSG00000178734	chr13: 75,871,038-75,883,811	0.83 (0.72-0.96)	-0.186	0.014
NPSR1-AS1	ENSG00000197085.7	chr7: 34,346,512-34,871,582	1.17 (1.04-1.31)	0.155	0.008
PRKG1-AS1	ENSG00000236671	chr10: 52,230,742-52,314,507	1.16 (1.13-1.31)	0.151	0.015

^aDerived from the multivariate Cox regression analysis of 264 lung adenocarcinoma patients in the training cohort. OGFRP1, opioid growth factor receptor pseudogene 1; ITGB1-DT, integrin subunit beta 1 divergent transcript; LMO7DN, LIM domain 7 downstream neighbor; NPSR1-AS1, neuropeptide S receptor 1-antisense RNA 1; PRKG1-AS1, protein kinase cyclic GMP-dependent 1-antisense RNA 1; chr, chromatin.

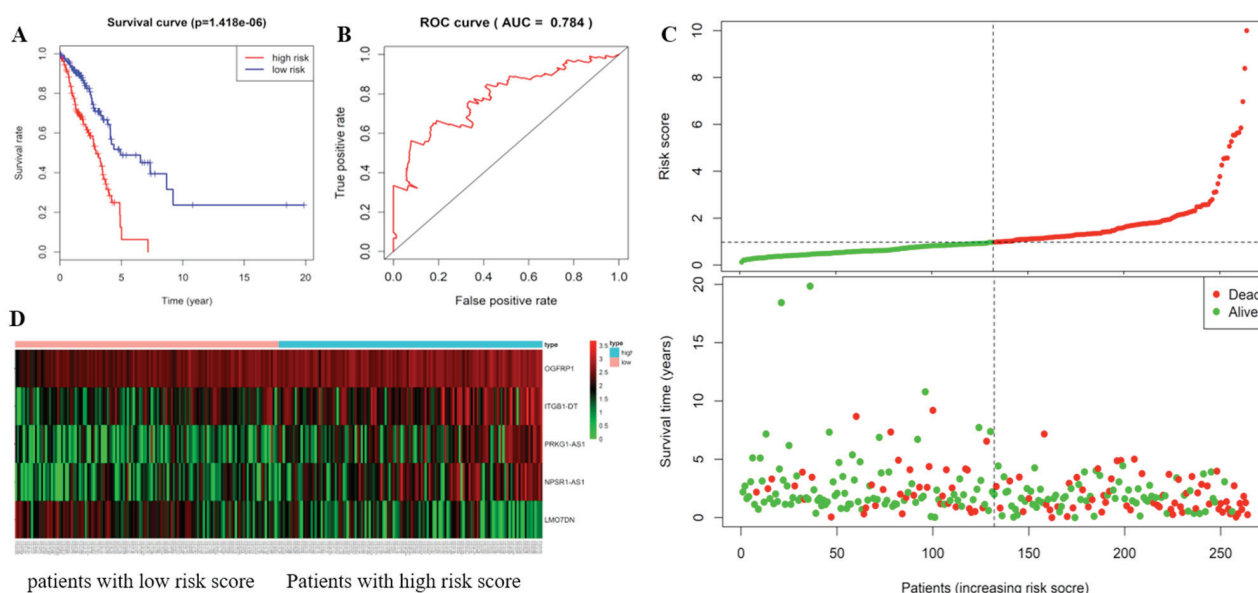


Figure 1. The 5 lncRNAs are able to predict overall survival in the training cohort. (A) Kaplan-Meier survival curves, (B) ROC curve for the training cohort, and the AUC value more than 0.7 is considered to be reliable. (C) risk distribution, survival status of patients and (D) heatmap of the 5 lncRNAs expression profiles in the training cohort. The '+' on the survival curves represents the censored data-points. lncRNA, long non-coding RNA; ROC, receiver operating characteristic; AUC, area under ROC curve; OGFRP1, opioid growth factor receptor pseudogene 1; ITGB1-DT, integrin subunit beta 1 divergent transcript; LMO7DN, LIM domain 7 downstream neighbor; NPSR1-AS1, neuropeptide S receptor 1-antisense RNA 1; PRKG1-AS1, protein kinase cyclic GMP-dependent 1-antisense RNA 1.

CI=1.636-3.027). More detailed results are listed in Table III. In short, the results indicated that the 5-lncRNA signature had good reliability and stability in predicting the survival of patients with LUAD.

The 5-lncRNA signature is an independent predictor of survival. Multivariate Cox regression analysis was performed on the 5-lncRNA signature in each cohort to assess whether the predictive ability of the prognostic signature was independent of other clinical factors, including sex age and TNM stage. In the multivariate Cox regression analysis, the OS was used as a dependent variable and the other clinical factors were regarded as covariates. The results indicated that the 5-lncRNA signature was significantly associated with the OS of patients with LUAD after adjustment by the other clinical factors in all cohorts (HR=2.418, 95% CI, 1.566-3.736, P<0.001 in the

training cohort; HR=1.925, 95% CI, 1.201-3.085, P=0.007 in the verification cohort; HR=2.117, 95% CI, 1.550-2.891, P<0.001 in the total cohort). More detailed results are listed in Table III. However, it was indicated that the TNM stage was also significantly associated with survival. Therefore, a data stratification analysis was performed to evaluate whether the prognostic signature still had a predictive value at the same TNM stage. All 486 patients with LUAD were stratified into three groups according to their TNM stage: The stage I (n=263), stage II (n=117) and stage III/IV (n=106) groups. With the median RS derived from the training cohort as a cut-off point, the stage I, II and III/IV groups were divided into high- (n=122, n=59 and n=64) and low-risk groups (n=141, n=58 and n=42), respectively. Significant differences in OS were observed between the high- and low-risk groups (log-rank test P=9.377x10⁻⁴, median survival: 1.67 vs. 2.08 years in the

Table III. Univariate and multivariate Cox regression analysis in each cohort.

A, Training cohort (n=264)						
Variables	Univariate analysis			Multivariate analysis		
	HR	95% CI of HR	P-value	HR	95% CI of HR	P-value
Five-lncRNA risk score (high/low risk)	2.743	1.792-4.200	<0.001	2.418	1.566-3.736	<0.001
Age (≥ 65 / < 65 years)	1.006	0.680-1.490	0.975	1.013	0.681-1.508	0.949
Sex (female/male)	1.078	0.724-1.603	0.712	1.026	0.682-1.542	0.903
Stage						
(II vs. I)	3.028	1.855-4.943	<0.001	2.583	1.568-4.253	<0.001
(III/IV vs. I)	2.837	1.774-4.537	<0.001	2.517	1.563-4.054	<0.001
B, Verification cohort (n=222)						
Variables	Univariate analysis			Multivariate analysis		
	HR	95% CI of HR	P-value	HR	95% CI of HR	P-value
Five-lncRNA risk score (high/low risk)	1.962	1.232-3.127	0.005	1.925	1.201-3.085	0.007
Age (≥ 65 / < 65 years)	1.369	0.860-2.181	0.185	1.442	0.886-2.283	0.145
Sex (female/male)	0.734	0.468-1.150	0.177	0.819	0.521-1.288	0.388
Stage						
(II vs. I)	1.655	0.944-2.902	0.079	1.692	0.948-3.020	0.075
(III/IV vs. I)	4.297	2.458-7.429	<0.001	4.515	2.571-7.929	<0.001
Total cohort (n=486)						
Variables	Univariate analysis			Multivariate analysis		
	HR	95% CI of HR	P-value	HR	95% CI of HR	P-value
Five-lncRNA risk score (high/low risk)	2.225	1.636-3.027	<0.001	2.117	1.550-2.891	<0.001
Age (≥ 65 / < 65 years)	1.147	0.851-1.547	0.368	1.198	0.887-1.619	0.239
Sex (female/male)	0.909	0.677-1.220	0.524	0.915	0.680-1.232	0.559
Stage						
(II vs. I)	2.318	1.606-3.347	<0.001	2.204	1.520-3.197	<0.001
(III/IV vs. I)	3.371	2.370-4.795	<0.001	3.284	2.301-4.687	<0.001

HR, hazard ratio; lncRNA, long non-coding RNA.

stage I group, Fig. 3A; log-rank test $P=8.543 \times 10^{-2}$; median survival: 1.73 vs. 1.90 years in the stage II group, Fig. 3B; log-rank test $P=5.078 \times 10^{-3}$; median survival: 1.25 vs. 1.72 years in the stage III/IV group, Fig. 3C). Although the P-value of the stage II group was above the significance level, a significant difference in median survival was observed between the high- and low-risk groups. These results demonstrated that the 5-lncRNA signature was an independent survival predictor for patients with LUAD.

Verification of the 5-lncRNA signature in GEO datasets. In order to determine the value of the 5-lncRNA signature in predicting survival, GEO datasets were analyzed. However,

when the 5 lncRNAs were searched in GEO datasets, the expression data of ITGB1-DT appeared not to be available. The 4 remaining lncRNAs were analyzed in a cohort composed of three GEO datasets (GSE30210, GSE30219 and GSE50081). As expected, the results indicated that OGFRL1, PRKG1-AS1 and LMO7DN were closely associated with the survival of patients with LUAD ($P<0.001$; Fig. 4). Among them, OGFRL1 and PRKG1-AS1 were risk factors for patients with LUAD, as their high expression was associated with poor survival, while LMO7DN acted as a protective factor, whose high expression was associated with a favorable outcome. Although there was no significant association between NPSR1-AS1 and survival in this cohort, the median survival in the high-expression

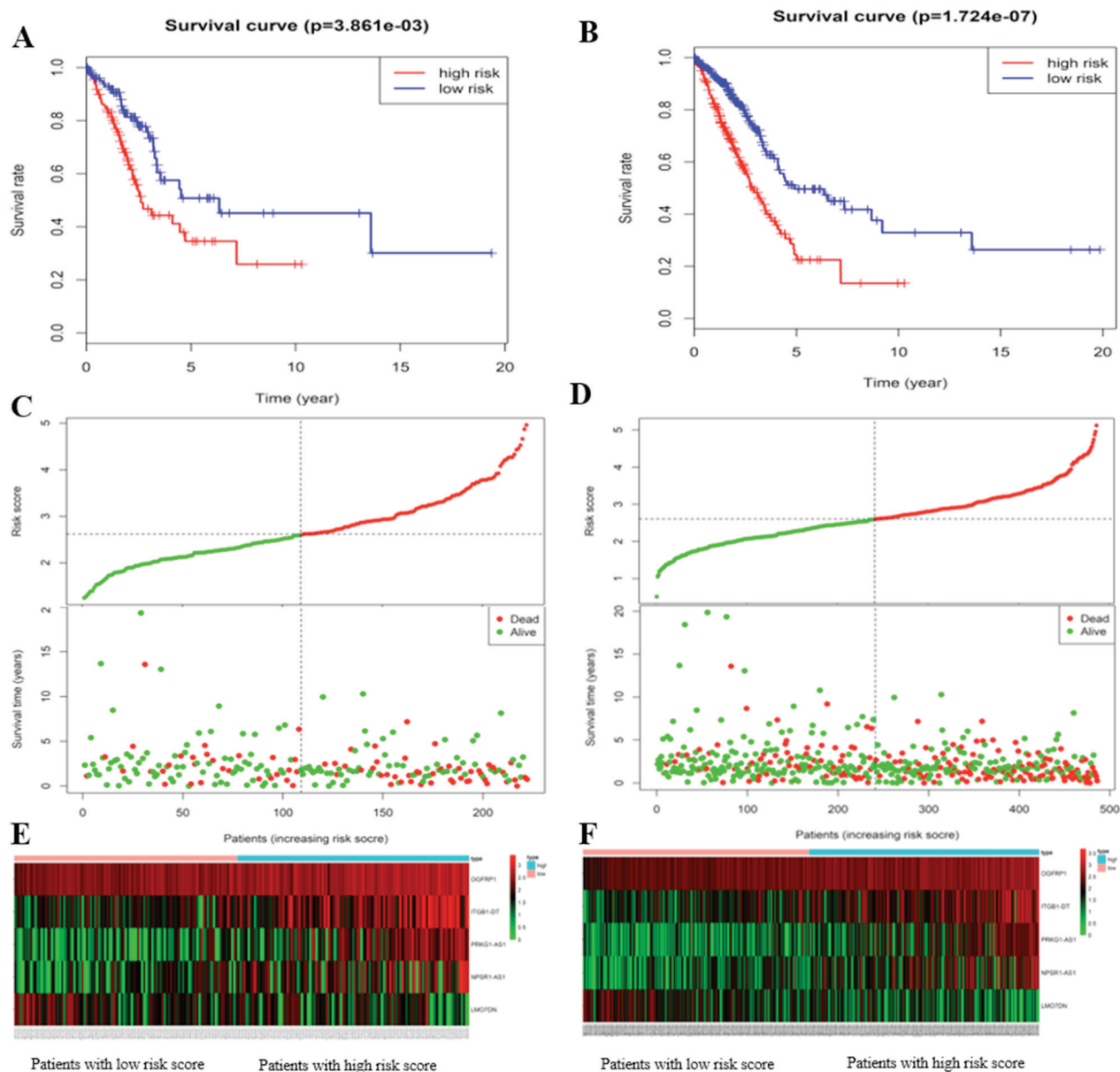


Figure 2. The 5 lncRNAs are able to predict the overall survival in the verifying and total cohorts. Kaplan-Meier survival curves for (A) verifying and (B) total cohorts, risk distribution, survival status of patients for (C) verifying and (D) total cohorts, and heat map of the 5 lncRNAs expression profiles in the (E) verifying and (F) total cohorts. The '+' on the survival curves represents the censored data-points. lncRNA, long non-coding RNA; OGFRP1, opioid growth factor receptor pseudogene 1; ITGB1-DT, integrin subunit beta 1 divergent transcript; LMO7DN, LIM domain 7 downstream neighbor; NPSR1-AS1, neuropeptide S receptor 1-antisense RNA 1; PRKG1-AS1, protein kinase cyclic GMP-dependent 1-antisense RNA 1.

group (4.54 years) was lower than that in the low-expression group (5.09 years) when the LUAD patients were divided into two groups with the median value of its expression used as a cut-off point. Based on the expression of the 4 lncRNAs in the GEO datasets, the reliability of the 5-lncRNA signature was further confirmed.

Role of ITGB1-DT in LUAD. Since the lncRNA ITGB1-DT was not present in the GEO datasets examined and no literature is currently available for it, it was assumed that this lncRNA has remained largely unexplored. Therefore, the role of ITGB1-DT in the tumorigenesis of LUAD was further investigated. Kaplan-Meier survival analysis for ITGB1-DT was performed, revealing that its expression levels are closely linked to the survival of patients with LUAD ($P=3.416 \times 10^{-4}$; Fig. 5A). Next,

the association between ITGB1-DT and the TNM stage, tumor stage and lymph node metastasis were analyzed, and the results indicated that ITGB1-DT was significantly associated with the tumor node metastasis (TNM stage; $P<0.005$; Fig. 5B), tumor stage ($P<0.005$; Fig. 5C) and lymph node metastasis ($P<0.05$; Fig. 5D). The LUAD patients with a higher TNM stage, tended to have a higher expression of ITGB1-DT.

Functional enrichment analysis for 5 prognostic lncRNAs. Pearson's correlation coefficients between the 5 lncRNAs and PCGs were calculated in the total cohort in order to further investigate the potential biological processes and pathways of these lncRNAs in LUAD. The results indicated that a total of 1,068 PCGs were closely associated with at least one of the 5 prognostic lncRNAs (Pearson's correlation coefficient) >0.40

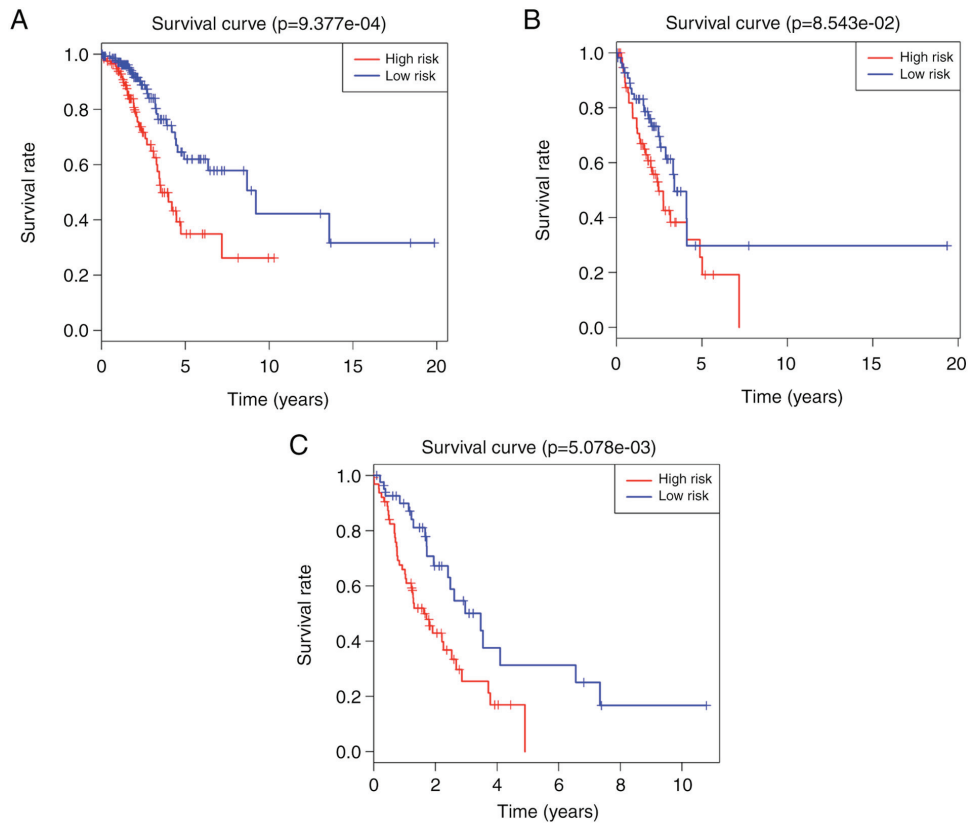


Figure 3. Survival analysis on the high- and low-risk groups in each tumor node metastasis stage, based on the 5-long non-coding RNA signature. Kaplan-Meier survival curves for (A) stage I, (B) stage II and (C) stage III/IV groups, and the '+' on the survival curves represents the censored data-points.

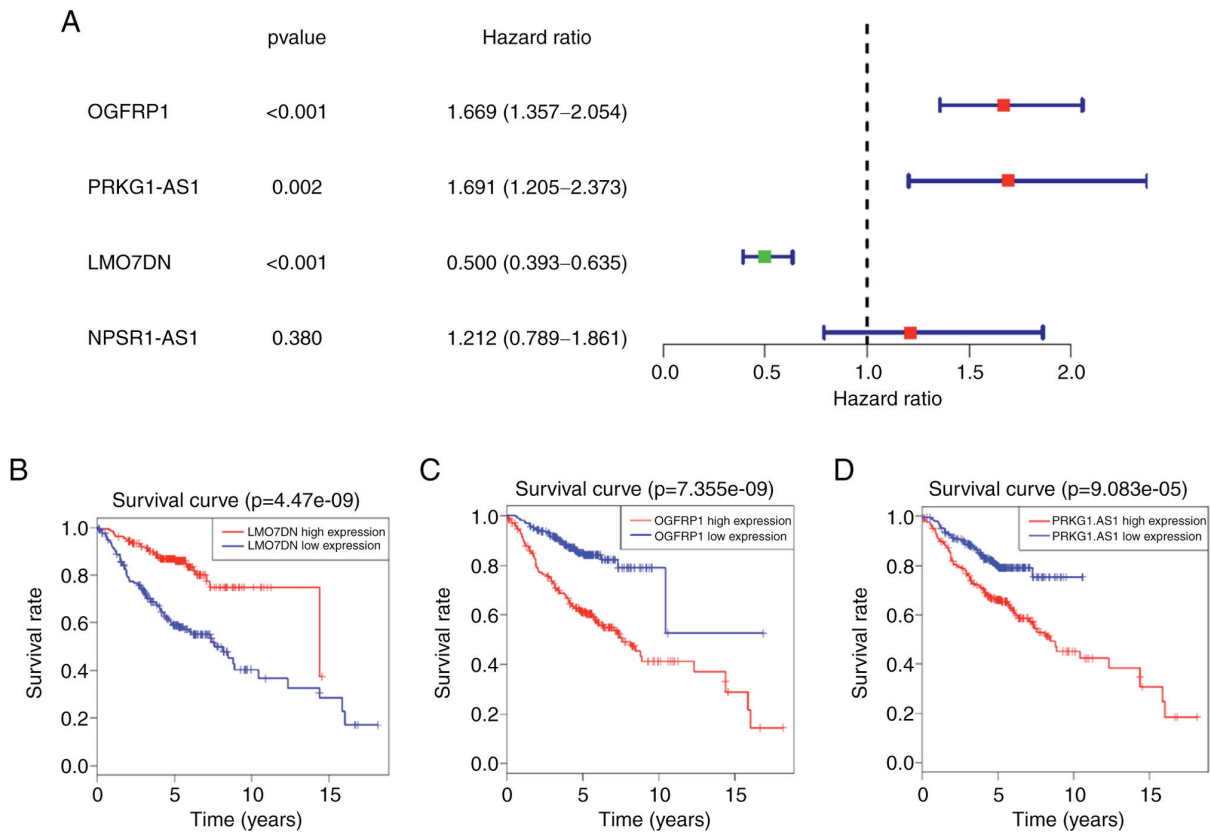


Figure 4. Analysis of GEO datasets for LMO7DN, NPSR1-AS1, OGFRP1 and PRKG1-AS1. (A) Results of the univariate Cox regression analysis of the 4 lncRNAs. Kaplan-Meier survival curves for (B) LMO7DN, (C) OGFRP1 and (D) PRKG1-AS1 in the GEO datasets, and the '+' on the survival curves represents the censored data-points. GEO, gene expression omnibus; lncRNA, long non-coding RNA; OGFRP1, opioid growth factor receptor pseudogene 1; LMO7DN, LIM domain 7 downstream neighbor; NPSR1-AS1, neuropeptide S receptor 1-antisense RNA 1; PRKG1-AS1, protein kinase cyclic GMP-dependent 1-antisense RNA 1.

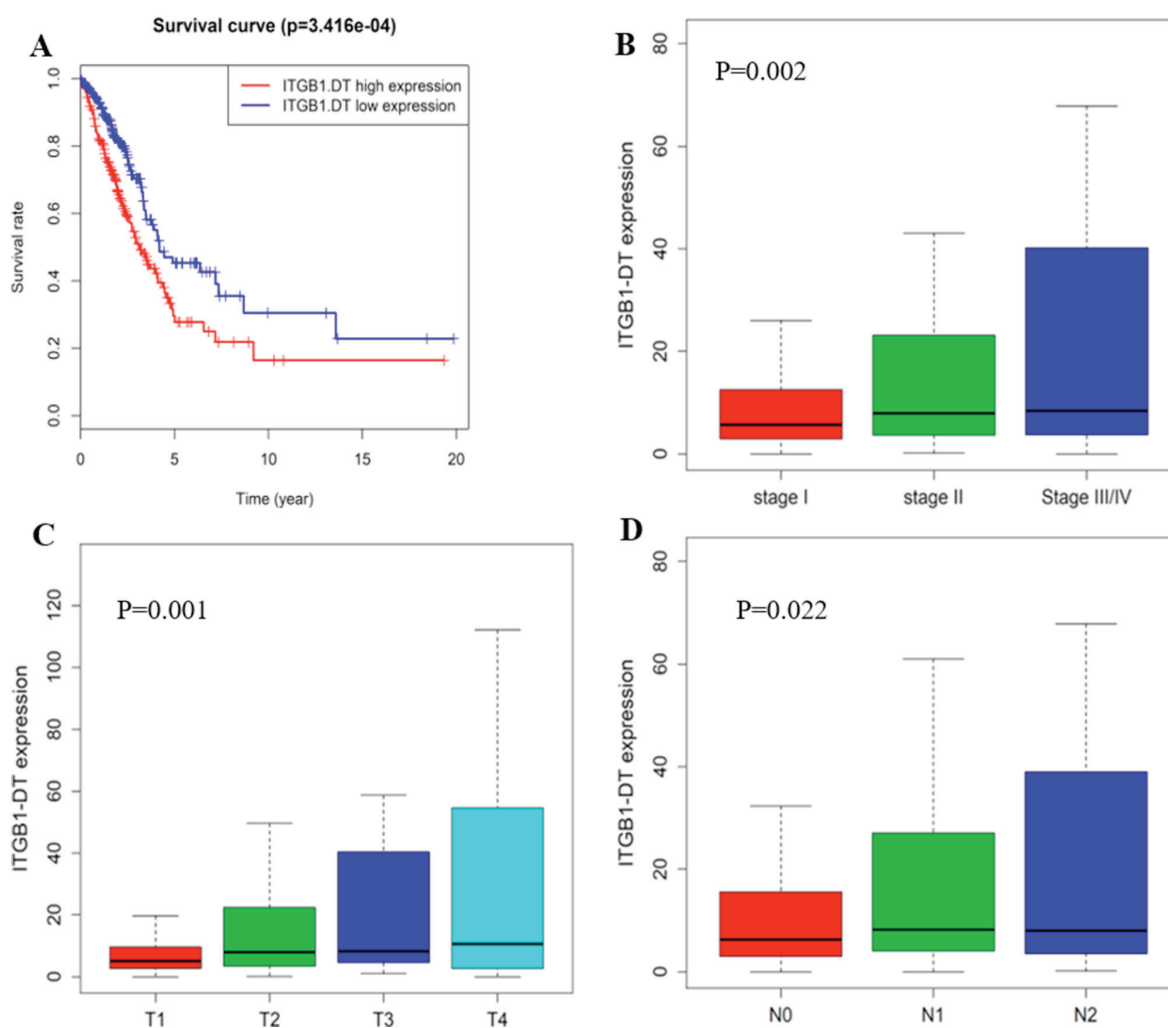


Figure 5. ITGB1-DT is significantly associated with the clinical factors in lung adenocarcinoma. (A) Kaplan-Meier survival curves for ITGB1-DT in the total cohort, and the '+' on the survival curves represents the censored data-points. (B-D) ITGB1-DT was significantly associated with (B) the TNM stage, (C) the tumor stage and (D) lymph node metastasis. Kruskal-Wallis test was used to compare the expression levels for ITGB1-DT across groups in each clinical factor. ITGB1-DT, integrin subunit beta 1 divergent transcript; TNM, tumor node metastasis.

and $P<0.01$). According to the functional enrichment analysis, the 1,068 PCGs were mainly enriched in 51 GO terms and 37 KEGG pathways ($P<0.005$). These GO terms were clustered into 5 major categories, including protein stability regulation, DNA replication and repair, signal regulation, cell division and cell cycle (Fig. 6A). A total of thirteen pathways were regarded as being most closely associated with tumorigenesis, including cell cycle, calcium signaling pathway, p53 signaling pathway, pathways in cancers, DNA replication, central carbon metabolism in cancer, homologous recombination, apoptosis and ras signaling pathway (Fig. 6B). Most of these GO terms and KEGG pathways have been indicated to be closely linked to the occurrence and development of LUAD, suggesting that the 5 lncRNAs may be involved in the LUAD-associated biological functions through co-expressed PCGs.

Expression levels of 5 lncRNAs in LUAD patients and cell lines.

The expression levels of the 5 lncRNAs of the signature were detected in 14 paired LUAD tissues and adjacent non-tumor tissues using RT-qPCR. Among these LUAD patients, six were female and eight were male; the average age was 58 years and the median age was 59 years (range, 40-80 years). The expression

levels of the 5 lncRNAs in lung cancer cell lines and a normal lung epithelial cell line were also detected by RT-qPCR. The expression levels of the 5 lncRNAs in LUAD tissues and cell lines were consistent with the results obtained with the TCGA dataset. The results revealed that ITGB1-DT, NPSR1-AS1, OGFRP1 and PRKG1-AS1 were significantly upregulated in LUAD tissues and cell lines, while LMO7DN was down-regulated in the LUAD tissues and cell lines compared with the normal control tissues and the normal cell line, respectively. The detailed information is provided in Fig. 7A-J.

Discussion

In recent years, a growing body of evidence has suggested that abnormal expression of lncRNAs is involved in various cancer-associated processes (36,37). In addition, certain lncRNAs have been indicated to be specific to tissues, disease types and developmental stages (38,39). An in-depth study of the association between lncRNAs and cancer is crucial for obtaining a better understanding of cancer. At present, although a large number of lncRNAs have been reported to have the potential to act as diagnostic, prognostic and therapeutic

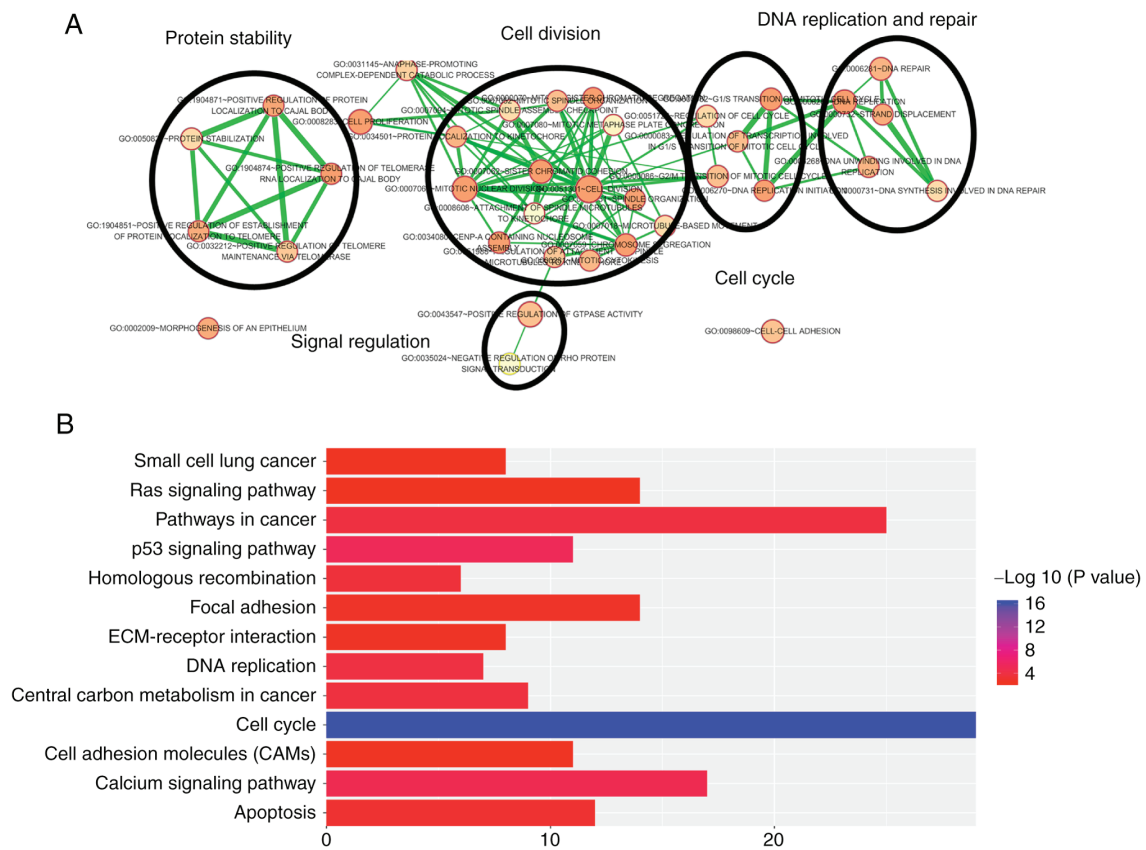


Figure 6. Functional enrichment analysis of 1,068 protein-coding genes co-expressed with the 5 prognostic long non-coding RNAs. (A) Functional enrichment map of GO terms. Each GO term is represented by a node and the overlapping genes between connecting terms are represented by an edge. The node size represents the number of genes in the GO terms. The color intensity represents the degree of enrichment. (B) Significantly enriched Kyoto Encyclopedia of Genes and Genomes pathways. GO, gene ontology; ECM, extracellular matrix.

targets, the clinical application of lncRNAs remains limited. In the present study, a 5-lncRNA signature was determined in order to provide a reference for identifying the ideal prognostic signature for clinical application.

In the present study, the top 7 lncRNAs were selected as candidate lncRNAs by analyzing the association between differentially expressed lncRNAs and the survival of LUAD patients using univariate Cox regression analysis in the training cohort. Next, multivariate Cox regression analysis, which used a model based on AIC, was performed on candidate lncRNAs. The model based on AIC was applied to construct a prognostic signature with the best predictive ability and the least number of lncRNAs. The prognostic signature that contained fewer genes is more likely to be applied in clinical practice. An RS model consisting of 5 lncRNAs, including OGFRP1, ITGB1-DT, LMO7DN, NPSR1-AS1 and PRKG1-AS1 was successfully established. The 5-lncRNA signature was proven to have good reliability and stability in predicting the prognosis of patients with LUAD through testing in the verification and total cohorts. Using multivariate Cox regression and data stratification analysis in each cohort, it was found that the 5-lncRNA signature could independently predict the prognosis of LUAD patients. The results suggested that the 5-lncRNA signature has a high potential to act as a prognostic biomarker.

Among the 5 lncRNAs, OGFRP1, NPSR1-AS1, ITGB1-DT and PRKG1-AS1 were indicated to be risk factors for the survival of patients with LUAD, while LMO7DN

was a protective factor. Of note, except for OGFRP1, these lncRNAs have remained largely unexplored and they were identified as prognostic cancer biomarkers for the first time in the present study, to the best of our knowledge. The expression of the 5 lncRNAs was then verified in LUAD tissues and cell lines. As expected, ITGB1-DT, NPSR1-AS1, OGFRP1 and PRKG1-AS1 were upregulated in LUAD tumor tissues and cell lines, while LMO7DN expression was downregulated compared with that in non-cancerous tissues and cell line. Of note, no expression data for ITGB1-DT were available in the GEO datasets examined. Further study on ITGB1-DT indicated that it was significantly associated with tumor size, lymph node metastasis and stage. In short, ITGB1-DT was indicated to have a key role in the occurrence and development of LUAD, and its in-depth study may provide novel insight into LUAD. A previous study suggested that OGFRP1 may be involved in the progression of hepatocellular carcinoma through the AKT/mTOR and Wnt/ β -catenin signaling pathways (40). Different studies demonstrated that OGFRP1 may influence the development of endometrial cancer (41) and NSCLC (42) by regulating miR-124-3p. OGFRP1 has also been identified as one of the lncRNA-associated signatures for predicting the survival of patients with LUAD (26), which further confirmed the potential value of the 5-lncRNA signature in predicting patient prognosis. Further research into the function of the other 4 lncRNAs may provide a better understanding of LUAD.

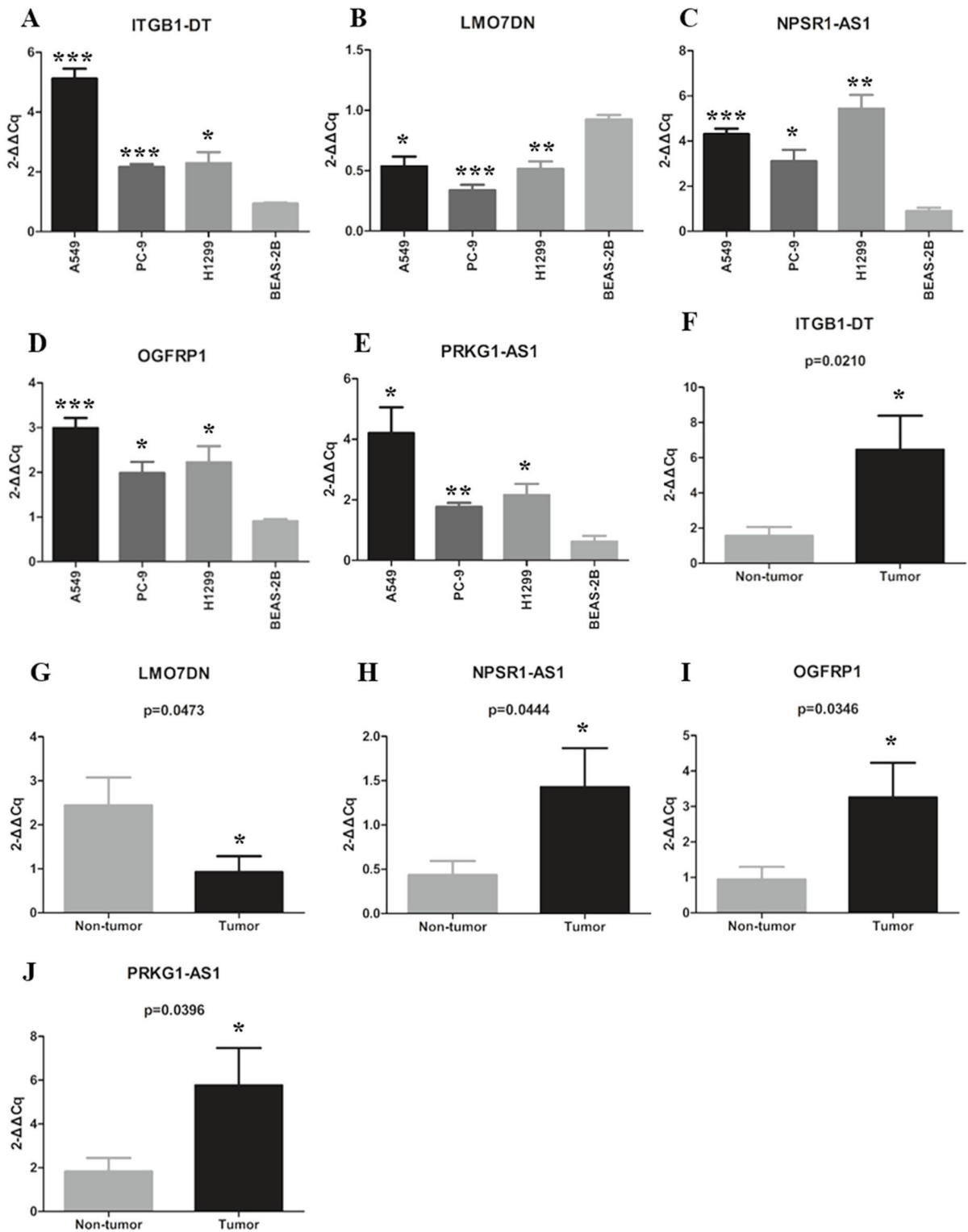


Figure 7. Expression of the 5 long non-coding RNAs in lung adenocarcinoma tissues and cell lines. (A-E) Differential expression of (A) ITGB1-DT, (B) LMO7DN, (C) NPSR1-AS1, (D) OGFRP1 and (E) PRKG1-AS1 between tumor and non-tumor tissues. (F-J) Differential expression of (F) ITGB1-DT, (G) LMO7DN, (H) NPSR1-AS1, (I) OGFRP1 and (J) PRKG1-AS1 between lung cancer cell lines and a normal human lung epithelial cell line. *P<0.05, **P<0.01, ***P<0.001 vs. non-tumor or BEAS-2B. OGFRP1, opioid growth factor receptor pseudogene 1; ITGB1-DT, integrin subunit beta 1 divergent transcript; LMO7DN, LIM domain 7 downstream neighbor; NPSR1-AS1, neuropeptide S receptor 1-antisense RNA 1; PRKG1-AS1, protein kinase cyclic GMP-dependent 1-antisense RNA 1.

To date, an increasing number of lncRNAs have been identified due to the development of various technologies (6). However, the great majority of lncRNAs, including the 5 lncRNAs identified in the present study, have not been well

characterized in terms of their functions. It has been reported that lncRNAs may be involved in the biological processes by interacting PCGs (43), which implied that the biological functions of lncRNAs may be predicted by analyzing the

co-expressed PCGs. In the present study, a total of 1,068 PCGs were considered to be closely associated with at least one of the 5 lncRNAs. A total of 5 functional categories and 13 KEGG pathways were obtained by performing functional enrichment analysis on these PCGs. These functional categories and KEGG pathways were all closely associated with tumorigenesis. For instance, the p53 signaling pathway has been indicated to be a key pathway in the process of tumorigenesis (44) and the cell cycle is a key biological process in tumorigenesis (45). The present results indicated that the 5 prognostic lncRNAs have an important role in LUAD via their involvement in these known cancer-associated biological functions.

In conclusion, a 5-lncRNAs signature with the ability to effectively predict the survival of LUAD patients was successfully established in the present study. The prognostic signature was proven to have good reliability and stability in predicting survival and maintain an independent predictive ability from other clinical factors. Furthermore, the 5 lncRNAs were indicated to involved in the tumorigenesis of LUAD through cancer-associated biological processes and pathways. Overall, the present results suggested that the 5-lncRNAs signature has the potential to act as an independent prognostic biomarker for LUAD and provide novel insight into the potential mechanisms of LUAD.

Acknowledgements

Not applicable.

Funding

The present study was funded by the National Natural Science Foundation of China (grant no. 81672300).

Availability of data and materials

The data used in this study were obtained from The Cancer Genome Atlas database (<https://portal.gdc.cancer.gov>) and Gene Expression Omnibus database (<http://www.ncbi.nlm.nih.gov/geo>).

Authors' contributions

LZ analyzed the data and wrote the manuscript. WW and YC were responsible for downloading the data. XL, JY and RS were responsible for selecting the literature. SY was responsible for the conception and experimental guidance of the study. All authors approved the final manuscript.

Ethics approval and consent to participate

Not applicable.

Patient consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Siegel RL, Miller KD and Jemal A: Cancer statistics, 2019. *CA Cancer J Clin* 69: 7-34, 2019.
2. Siegel RL, Miller KD and Jemal A: Cancer statistics, 2017. *CA Cancer J Clin* 67: 7-30, 2017.
3. Sholl LM: The molecular pathology of lung cancer. *Surg Pathol Clin* 9: 353-378, 2016.
4. Mao Y, Yang D, He J and Krasna MJ: Epidemiology of lung cancer. *Surg Oncol Clin N Am* 25: 439-445, 2016.
5. Puri T: Targeted therapy in nonsmall cell lung cancer. *Indian J Cancer* 54: 83-88, 2017.
6. Jathar S, Kumar V, Srivastava J and Tripathi V: Technological developments in lncRNA biology. *Adv Exp Med Biol* 1008: 283-323, 2017.
7. Jarroux J, Morillon A and Pinskaya M: History, discovery, and classification of lncRNAs. *Adv Exp Med Biol* 1008: 1-46, 2017.
8. Bhan A, Soleimani M and Mandal SS: Long noncoding RNA and cancer: A New Paradigm. *Cancer Res* 77: 3965-3981, 2017.
9. Liu Y, Sharma S and Watabe K: Roles of lncRNA in breast cancer. *Front Biosci (Schol Ed)* 7: 94-108, 2015.
10. Thin KZ, Liu X, Feng X, Raveendran S and Tu JC: LncRNA-DANCR: A valuable cancer related long non-coding RNA for human cancers. *Pathol Res Pract* 214: 801-805, 2018.
11. Gibb EA, Brown CJ and Lam WL: The functional role of long non-coding RNA in human carcinomas. *Mol Cancer* 10: 38, 2011.
12. Balas MM and Johnson AM: Exploring the mechanisms behind long noncoding RNAs and cancer. *Noncoding RNA Res* 3: 108-117, 2018.
13. Qi G, Kong W, Mou X and Wang S: A new method for excavating feature lncRNA in lung adenocarcinoma based on pathway crosstalk analysis. *J Cell Biochem* 120: 9034-9046, 2019.
14. Peng W, Wang J, Shan B, Peng Z, Dong Y, Shi W, He D, Cheng Y, Zhao W, Zhang C, *et al*: Diagnostic and prognostic potential of circulating long non-coding RNAs in non small cell lung cancer. *Cell Physiol Biochem* 49: 816-827, 2018.
15. Dai SP, Jin J and Li WM: Diagnostic efficacy of long non-coding RNA in lung cancer: A systematic review and meta-analysis. *Postgrad Med J* 94: 578-587, 2018.
16. Pan X, Zheng G and Gao C: LncRNA PVT1: A novel therapeutic target for cancers. *Clin Lab* 64: 655-662, 2018.
17. Song P, Jiang B, Liu Z, Ding J, Liu S and Guan W: A three-lncRNA expression signature associated with the prognosis of gastric cancer patients. *Cancer Med* 6: 1154-1164, 2017.
18. Bao Z, Zhang W and Dong D: A potential prognostic lncRNA signature for predicting survival in patients with bladder urothelial carcinoma. *Oncotarget* 8: 10485-10497, 2017.
19. Gu JX, Zhang X, Miao RC, Xiang XH, Fu YN, Zhang JY, Liu C and Qu K: Six-long non-coding RNA signature predicts recurrence-free survival in hepatocellular carcinoma. *World J Gastroenterol* 25: 220-232, 2019.
20. Luo D, Deng B, Weng M, Luo Z and Nie X: A prognostic 4-lncRNA expression signature for lung squamous cell carcinoma. *Artif Cells Nanomed Biotechnol* 46: 1207-1214, 2018.
21. Lu T, Wang Y, Chen D, Liu J and Jiao W: Potential clinical application of lncRNAs in non-small cell lung cancer. *Onco Targets Ther* 11: 8045-8052, 2018.
22. Miao R, Ge C, Zhang X, He Y, Ma X, Xiang X, Gu J, Fu Y, Qu K, Liu C, *et al*: Combined eight-long noncoding RNA signature: A new risk score predicting prognosis in elderly non-small cell lung cancer patients. *Aging (Albany NY)* 11: 467-479, 2019.
23. Li X, Li B, Ran P and Wang L: Identification of ceRNA network based on a RNA-seq shows prognostic lncRNA biomarkers in human lung adenocarcinoma. *Oncol Lett* 16: 5697-5708, 2018.
24. Kumar P, Khadirnaikar S and Shukla SK: A novel lncRNA-based prognostic score reveals TP53-dependent subtype of lung adenocarcinoma with poor survival. *J Cell Physiol*: Feb 10, 2019 doi: 10.1002/jcp.28260 (Epub ahead of print).
25. Songyang Y, Zhu W, Liu C, Li LL, Hu W, Zhou Q, Zhang H, Li W and Li D: Large-scale gene expression analysis reveals robust gene signatures for prognosis prediction in lung adenocarcinoma. *PeerJ* 7: e6980, 2019.
26. Li YY, Yang C, Zhou P, Zhang S, Yao Y and Li D: Genome-scale analysis to identify prognostic markers and predict the survival of lung adenocarcinoma. *J Cell Biochem* 119: 8909-8921, 2018.
27. Robinson MD, McCarthy DJ and Smyth GK: edgeR: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26: 139-140, 2010.

28. Huang PH: Asymptotics of AIC, BIC, and RMSEA for model selection in structural equation modeling. *Psychometrika* 82: 407-426, 2017.
29. Sing T, Sander O, Beerenwinkel N and Lengauer T: ROCr: Visualizing classifier performance in R. *Bioinformatics* 21: 3940-3941, 2005.
30. Huang DW, Sherman BT, Tan Q, Kir J, Liu D, Bryant D, Guo Y, Stephens R, Baseler MW, Lane HC and Lempicki RA: DAVID bioinformatics resources: Expanded annotation database and novel algorithms to better extract biology from large gene lists. *Nucleic Acids Res* 35: W169-W175, 2007.
31. Wu J, Mao X, Cai T, Luo J and Wei L: KOBAS server: A web-based platform for automated annotation and pathway identification. *Nucleic Acids Res* 34: W720-W724, 2006.
32. Merico D, Isserlin R, Stueker O, Emili A and Bader GD: Enrichment map: A network-based method for gene-set enrichment visualization and interpretation. *PLoS One* 5: e13984, 2010.
33. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T: Cytoscape: A software environment for integrated models of biomolecular interaction networks. *Genome Res* 13: 2498-2504, 2003.
34. Ito K and Murphy D: Application of ggplot2 to pharmacometric graphics. *CPT Pharmacometrics Syst Pharmacol* 2: e79, 2013.
35. Livak KJ and Schmittgen TD: Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) method. *Methods* 25: 402-408, 2001.
36. Forrest ME and Khalil AM: Review: Regulation of the cancer epigenome by long non-coding RNAs. *Cancer Lett* 407: 106-112, 2017.
37. Yang G, Lu X and Yuan L: LncRNA: A link between RNA and cancer. *Biochim Biophys Acta* 1839: 1097-1109, 2014.
38. Schmitz SU, Grote P and Herrmann BG: Mechanisms of long noncoding RNA function in development and disease. *Cell Mol Life Sci* 73: 2491-2509, 2016.
39. Zhang H, Chen Z, Wang X, Huang Z, He Z and Chen Y: Long non-coding RNA: A new player in cancer. *J Hematol Oncol* 6: 37, 2013.
40. Chen W, You J, Zheng Q and Zhu YY: Downregulation of lncRNA OGFRP1 inhibits hepatocellular carcinoma progression by AKT/mTOR and Wnt/ β -catenin signaling pathways. *Cancer Manag Res* 10: 1817-1826, 2018.
41. Lv Y, Chen S, Wu J, Lin R, Zhou L, Chen G, Chen H and Ke Y: Upregulation of long non-coding RNA OGFRP1 facilitates endometrial cancer by regulating miR-124-3p/SIRT1 axis and by activating PI3K/AKT/GSK-3 β pathway. *Artif Cells Nanomed Biotechnol* 47: 2083-2090, 2019.
42. Tang LX, Chen GH, Li H, He P, Zhang Y and Xu XW: Long non-coding RNA OGFRP1 regulates LYPD3 expression by sponging miR-124-3p and promotes non-small cell lung cancer progression. *Biochem Biophys Res Commun* 505: 578-585, 2018.
43. Ferré F, Colantoni A and Helmer-Citterich M: Revealing protein-lncRNA interaction. *Brief Bioinform* 17: 106-116, 2016.
44. Joerger AC and Fersht AR: The p53 pathway: Origins, inactivation in cancer, and emerging therapeutic approaches. *Annu Rev Biochem* 85: 375-404, 2016.
45. Williams GH and Stoerber K: The cell cycle and cancer. *J Pathol* 226: 352-364, 2012.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) License.