

# An 11-lncRNA risk scoring model predicts prognosis of lung squamous cell carcinoma

J.-H. LIU<sup>1</sup>, Y.-M. XU<sup>2</sup>, J. YAN<sup>3</sup>

<sup>1</sup>Department of Respiration, Zhejiang Jinhua Guangfu Hospital, Jinhua, China

<sup>2</sup>Department of Respiratory Medicine, Shaoxing Municipal Hospital, Shaoxing, China

<sup>3</sup>Department of Respiratory, the 903<sup>rd</sup> Hospital of PLA, Hangzhou, China

**Abstract. – OBJECTIVE:** The study aims to construct a multi-gene risk scoring model that can be used to predict the prognosis of patients with lung squamous cell carcinoma (LUSC).

**PATIENTS AND METHODS:** RNA-seq data from 494 LUSC tumor samples and 49 peripheral lung tissue samples were obtained from TCGA\_LUSC database. Differential analysis was conducted using edgeR to screen differentially expressed lncRNAs (DElncRNAs) between LUSC and normal samples. Univariate Cox regression analysis was used to screen lncRNAs that were significantly correlated with LUSC prognosis. LASSO regression model was built to reduce complexity. The LUSC prognostic model based on lncRNAs was established by multivariate Cox regression analysis, which was evaluated by ROC curves and survival analysis. ROC and Kaplan-Meier survival curves of each lncRNA in the model were plotted and compared.

**RESULTS:** 2085 DElncRNAs were identified. Combined with univariate Cox regression analysis, 342 prognosis-related genes were screened. After LASSO regression analysis, 11 lncRNAs tightly related to LUSC prognosis were identified and a risk scoring model was constructed. ROC curve analysis proved the good performance of the model. The Kaplan-Meier survival curve showed that the mortality in high-risk group was significantly higher. The survival analysis results of each lncRNA were also consistent with the prediction in Cox regression.

**CONCLUSIONS:** Our results suggested that the 11-lncRNA risk scoring model may provide a new insight for predicting prognosis of LUSC patients.

*Key Words:*

lncRNA, Lung squamous cell carcinoma, Prognosis, Risk scoring.

## Introduction

Lung cancer is one of the most common malignancies in the world, and can be roughly divided into small cell lung cancer (SCLC) and non-small

cell lung cancer (NSCLC)<sup>1</sup>. Lung squamous cell carcinoma (LUSC) is the most common subtype of NSCLC with high molecular complexity<sup>2</sup>. It accounts for about 30% of all NSCLC<sup>3</sup>. Most LUSC patients were diagnosed with advanced stage<sup>4</sup>, resulting in poor therapeutic effect, high recurrence rate and poor prognosis<sup>5</sup>. The main reason is that LUSC cells often penetrate into adjacent tissues without a defined range. In some ways it contributes to a high level of LUSC recurrence, making survival outcomes more difficult to assess. In conclusion, effective prognosis prediction of LUSC is essential to improve the survival rate of LUSC patients<sup>6</sup>. It is urgent to find specific biomarkers which are deficient currently for LUSC prognosis in clinical practice<sup>7</sup>.

Long non-coding RNAs (lncRNAs) are a group of non-coding RNAs with a length of more than 200 nucleotides<sup>7</sup>. They are transcribed from thousands of sites in mammalian genome<sup>8</sup> and are important regulators in various biological and pathological processes<sup>9</sup>. In nucleus, lncRNAs are involved in the organization of chromatin, transcription and post-transcriptional gene expression regulation and other biological processes, which also act as the structural framework of nuclear domain<sup>10</sup>. In cytoplasm, lncRNA binds to ribosomes to regulate mRNA stability and translation<sup>11,12</sup>. lncRNAs may play a crucial role in the physiology and pathology of human diseases as oncogenes or tumor suppressor genes<sup>7</sup>. They can drive the initiation, proliferation and metastasis of tumors, and can be used as biomarkers for the diagnosis and prognosis of lung cancer<sup>13-16</sup>. So far, only a few lncRNA-based models for LUSC prognosis prediction have been developed<sup>6,17,18</sup> with different results. Thus, it is of great significance to study the prognostic value of lncRNAs in LUSC.

In this study, edgeR was used to analyze HTSeq-Counts data downloaded from the TC-

GA database for screening prognostic lncRNAs. Through the construction of LASSO model and Cox regression analysis, the 11-lncRNA risk scoring model related to the prognosis of LUSC was determined and provided a new reference for evaluating the prognosis of LUSC patients.

## Patients and Methods

### Data and Patients

LUSC and normal adjacent tissues were downloaded from The Cancer Genome Atlas (TCGA, <https://cancergenome.nih.gov/>), which included 494 LUSC and 49 unpaired normal lung tissue samples with corresponding clinical pathologic information. This study conforms to the publication guidelines provided by TCGA.

### Identification of Potential Prognostic lncRNAs Associated with Overall Survival (OS) in LUSC Patients

9735 lncRNAs expression profiles were obtained from the TCGA-LUSC group. EdgeR package ([bioconductor.org/packages/release/bioc/html/edgeR.html](http://bioconductor.org/packages/release/bioc/html/edgeR.html)) was used to identify differentially expressed lncRNAs (DELncRNAs) between LUSC and normal samples.  $|\log_{2}FC| > 2$  and  $p < 0.05$  were set as threshold. Univariate Cox regression analysis was used to evaluate each gene and to screen the lncRNAs that were significantly related to LUSC prognosis. In order to avoid over-fitting of the subsequent multivariate Cox regression model, the LASSO regression model was constructed to reduce the complexity.

Then, multivariate Cox regression was performed to obtain the risk scoring model related to LUSC prognosis, and the patients were ranked according to the risk score. In addition, we evaluated the predictive value through the time-dependent receiver operating characteristic (ROC) curves. The area under the ROC curves (AUC) was calculated to determine the performance of the risk scoring model. The Kaplan-Meier survival curve was used to compare survival differences between high-risk and low-risk groups. Finally, ROC and Kaplan-Meier survival curves of each lncRNA in the risk scoring model were compared with the those of the risk model.

### Statistical Analysis

Univariate and multivariate Cox regression analysis were carried out in TCGA dataset. LASSO statistical algorithm was carried out using the

“glmnet” software package (<https://www.r-project.org/>) in R software. Statistical analysis was performed using Statistical Product and Service Solution (SPSS) 23.0 (IBM Corp., Armonk, NY, USA), and  $p < 0.05$  was considered statistically significant unless otherwise indicated. GraphPad Prism5 (GraphPad Software Inc, San Diego, CA, USA) was used to plot ROC curves.

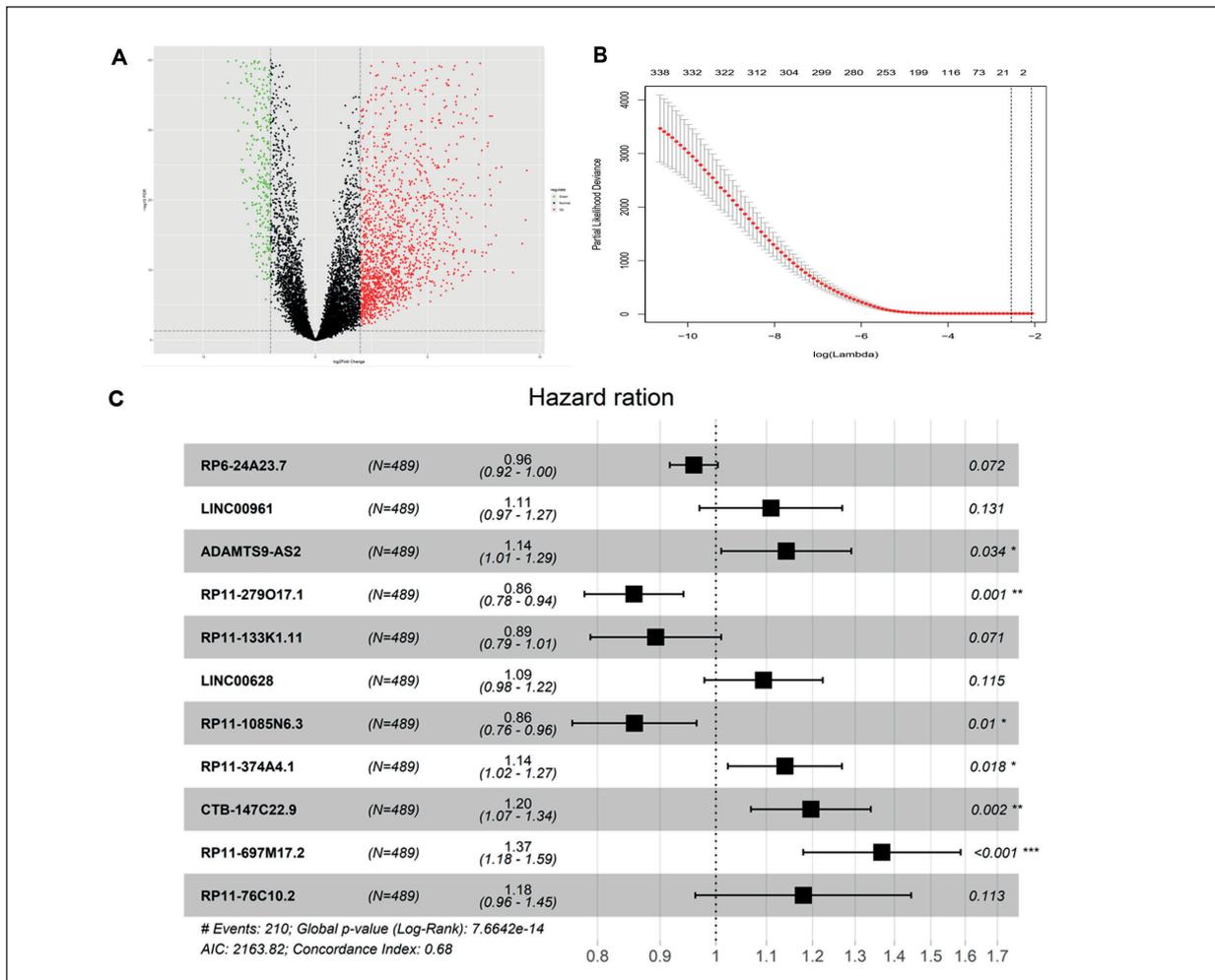
## Results

### Differential Expression Profiles and Identification of Prognostic lncRNAs

2085 DELncRNAs were screened out from the TCGA-LUSC database (Figure 1A). Then, univariate Cox regression analysis was performed, and 342 lncRNAs were screened with  $p < 0.1$  as the threshold. In order to prevent the over-fitting of the subsequent multivariate Cox regression model, LASSO regression model was constructed to further analyze the 342 genes (Figure 1B) to reduce the complexity. When the  $\beta$  value was 11, the simulation fitting was optimal. So, 11 genes were selected for model construction. Multivariate Cox proportional hazard regression analysis was performed on the 11 lncRNAs to further evaluate the prognostic value. As shown in Figure 1C, ADAMTS9-AS2, LINC00628, LINC00961, RP11-374A4.1, CTB-147C22.9, RP11-697M17.2, and RP11-76C10.2 were high-risk genes, while RP6-24A23.7, RP11-279O17.1, RP11-133K1.11 and RP11-1085N6.3 were low-risk genes.

### Assessment of the 11-lncRNA Risk Scoring Model in LUSC Patients from TCGA

According to the 11 lncRNAs screened by the above analysis, the risk scoring model was constructed as follows: Risk Score =  $(-0.041629506 \times RP6-24A23.7) + (0.10370246 \times LINC00961) + (0.13273787 \times ADAMTS9-AS2) + (-0.154425544 \times RP11-279O17.1) + (-0.113445439 \times RP11-133K1.11) + (0.089693068 \times LINC00628) + (-0.153722976 \times RP11-1085N6.3) + (0.130026784 \times RP11-374A4.1) + (0.178922776 \times CTB-147C22.9) + (0.312789492 \times RP11-697M17.2) + (0.164686718 \times RP11-76C10.2)$ . In order to identify the correlation between the risk score of the 11-lncRNA model and their clinicopathological characteristics, the patients from TCGA were divided into high-risk group and low-risk group. The median risk score was set as the critical value. As shown in Table I, there was no statistical difference in gender, age, smoking



**Figure 1.** Differential expression profiles and identification of prognostic lncRNAs. **A**, Volcano maps of DE lncRNAs (The red dots represent the up-regulated lncRNAs and the green dots represent the down-regulated lncRNAs). **B**, LASSO regression model. **C**, Multivariate Cox proportional hazard regression analysis of 11 lncRNAs.

status, and pathological stages between patients with a high-risk score and a low risk score, suggesting that the prognostic ability of the model was independent. The event of the high- and low-risk groups showed significant differences, indicating that the model could distinguish the survival status. In addition, the ROC curves showed that the AUC of the 11-lncRNA risk scoring model was 0.756 (Figure 2A). While the AUC of each lncRNA in this model (Figure 3) was smaller than that of the risk scoring model, indicating that risk scoring model was a good indicator of survival prediction with better accuracy. The Kaplan-Meier curve (Figure 2B) showed that the OS of the high-risk group was significantly shorter than that of the low-risk group ( $p < 0.05$ ).

Subsequently, the risk scores of the patients in the training group were ranked, and the survival

status of each patient was dot plotted (Figure 2C). It was found that the mortality rate of patients in the high-risk group was significantly higher than that of patients in the low-risk group. The expression heat map for lncRNAs (Figure 2D) showed that the expression levels of RP6-24A23.7, RP11-279O17.1, RP11-133K1.11 and RP11-1085N6.3 were decreased with the increase of risk score. While the expression levels of ADAMTS9-AS2, LINC00628, LINC00961, RP11-374A4.1, CTB-147C22.9, RP11-697M17.2 and RP11-76C10.2 were increased with the increase of risk score. It was consistent with the prediction result of multivariate Cox regression analysis.

### Survival Analysis Verification

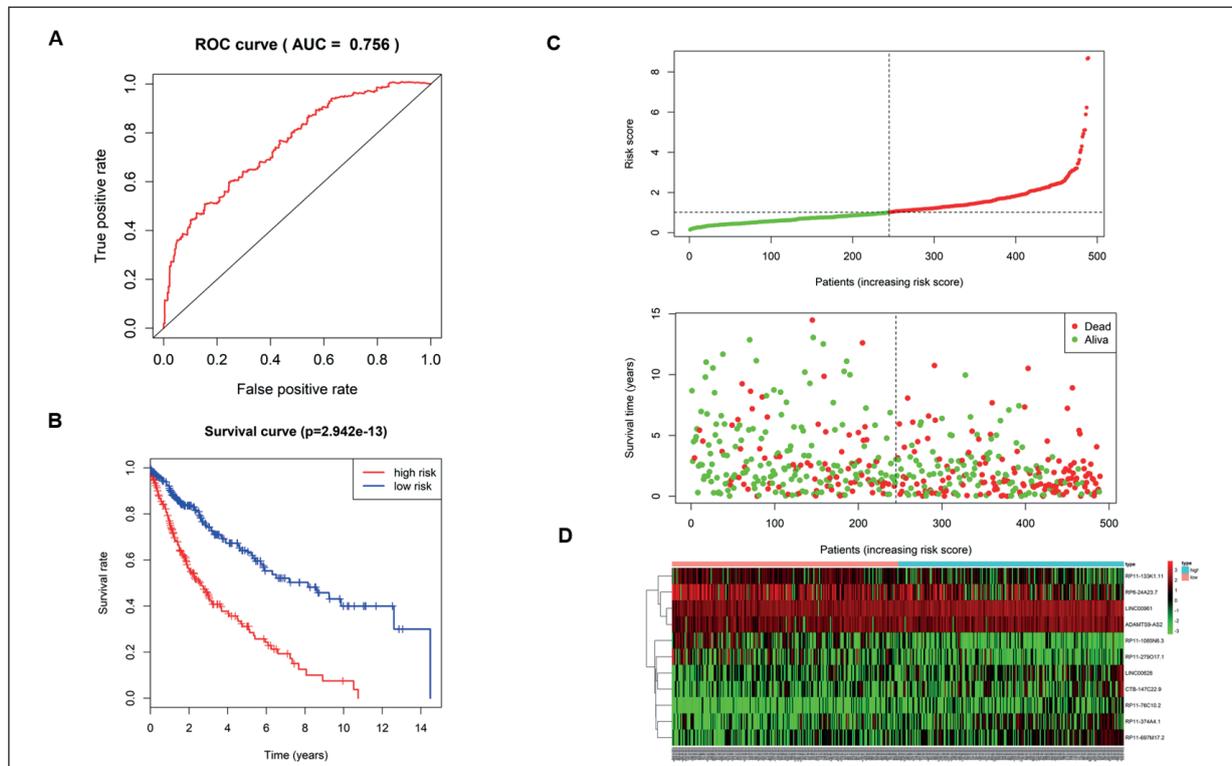
In the training set, Kaplan-Meier survival analysis was performed on the above 11 lncRNAs. The

**Table I.** Association between risk scores of 11-lncRNA markers and clinicopathological characteristics.

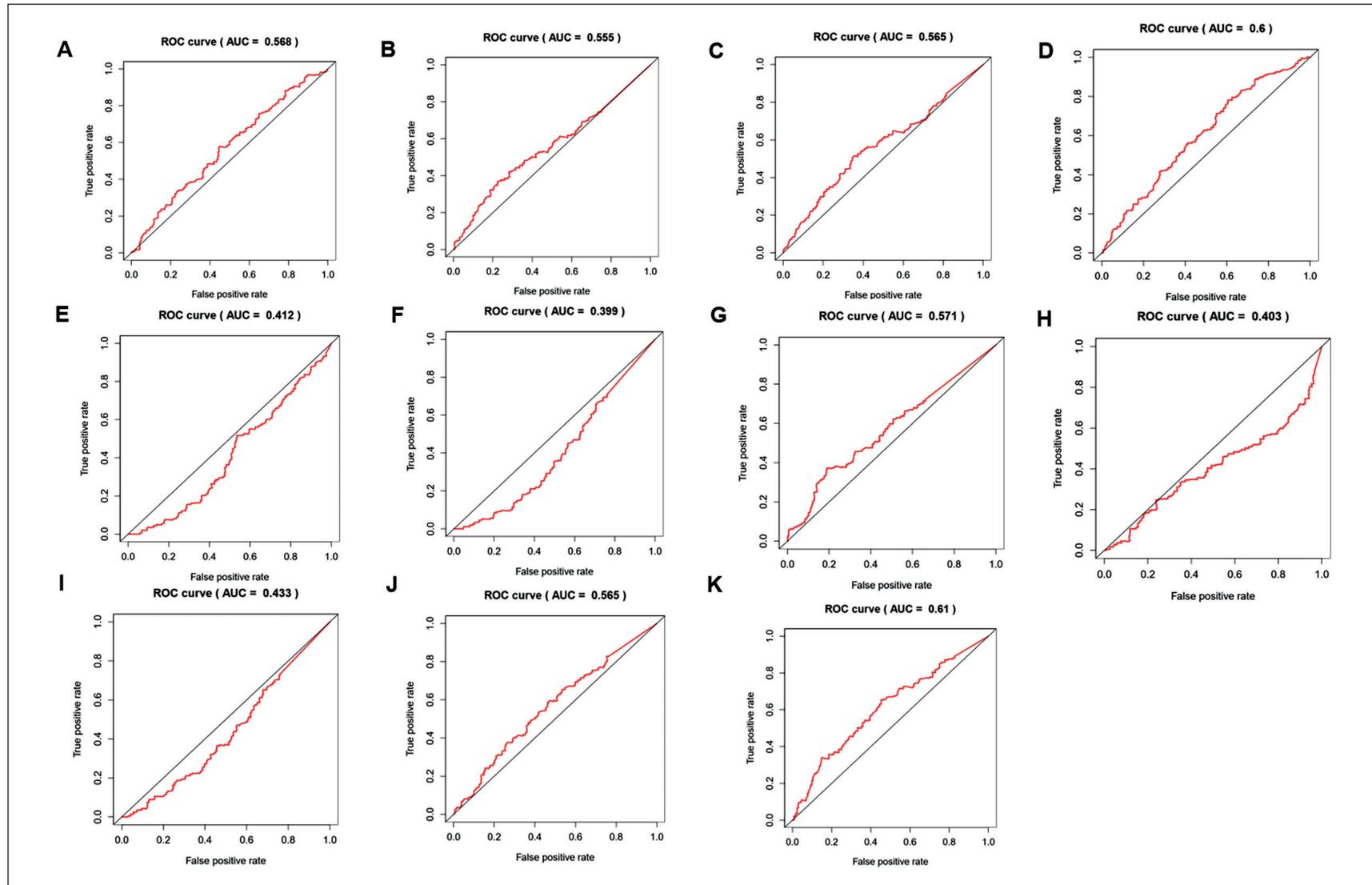
Variables	Low risk (n = 234)	High risk (n = 234)	p-value
Gender			0.116
Female	54 (23.1%)	70 (29.9%)	
Male	180 (76.9%)	164 (70.1%)	
Age (years)		0.463	
Age < 60	44 (18.8%)	37 (15.8%)	
Age ≥ 60	190 (81.2%)	197 (84.2%)	
Smoking (years)		0.603	
1	7 (3.0%)	11 (4.7%)	
2	62 (26.5%)	69 (29.5%)	
3	38 (16.2%)	43 (18.4%)	
4	124 (53.0%)	109 (46.6%)	
5	3 (1.3%)	2 (0.9%)	
Event			< 0.001
Yes	73 (31.2%)	132 (56.4%)	
No	161 (68.8%)	102 (43.6%)	
p stage			0.713
Stage I	116 (49.6%)	117 (50.0%)	
Stage II	75 (32.1%)	73 (31.2%)	
Stage III	41 (17.5%)	39 (16.7%)	
Stage IV	2 (0.9%)	5 (2.1%)	

results showed that patients with high expression of lncRNAs, including ADAMTS9-AS2 (Figure 4A), LINC00628 (Figure 4B), LINC00961

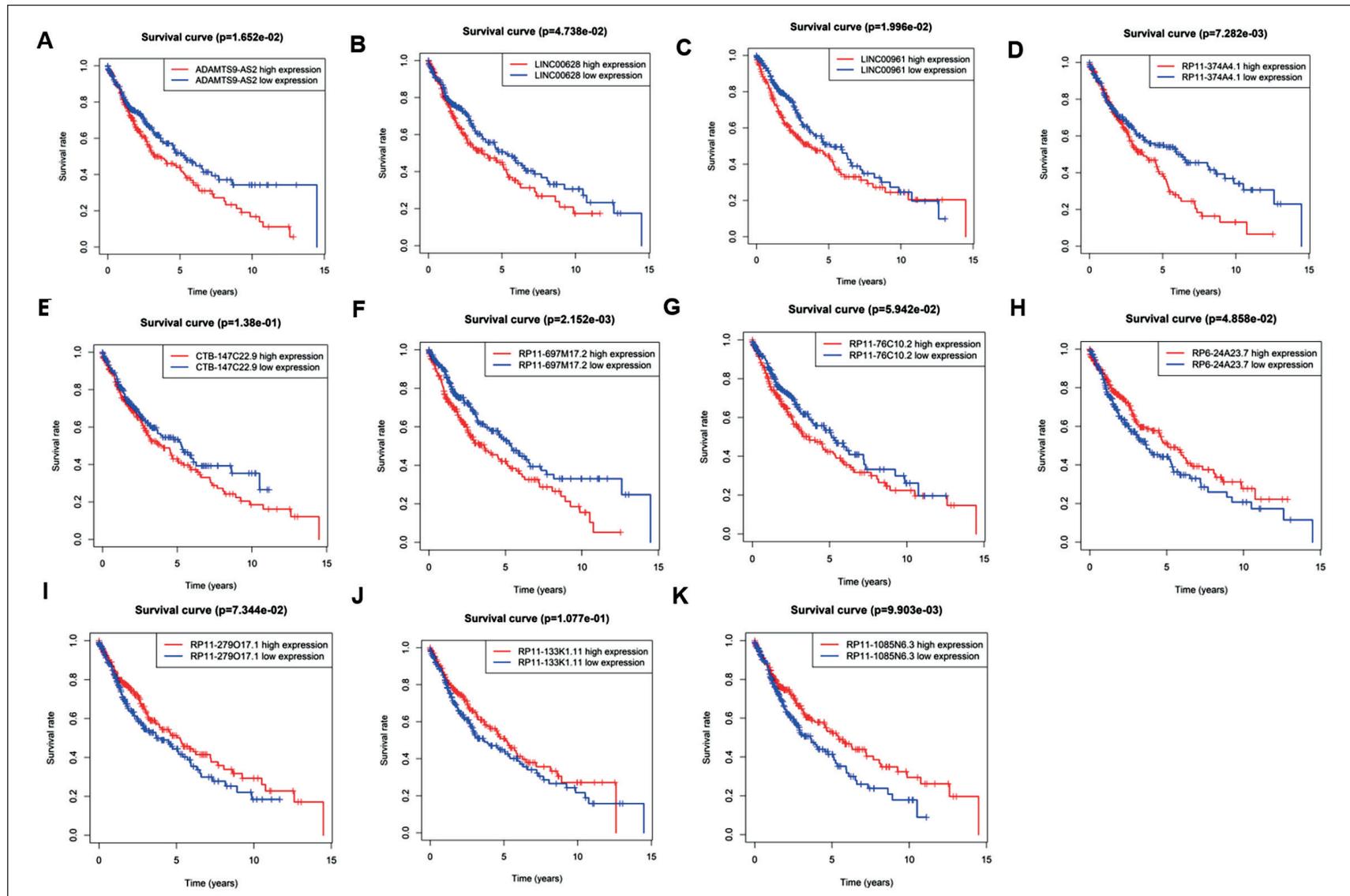
(Figure 4C), RP11-374A4.1 (Figure 4D), CTB-147C22.9 (Figure 4E), RP11-697M17.2 (Figure 4F) and RP11-76C10.2 (Figure 4G) tended to have



**Figure 2.** Assessment of the 11-lncRNA risk scoring model in patients from TCGA. **A**, ROC curve of the 11-lncRNA risk scoring model. **B**, Kaplan-Meier curve of the high-risk and low-risk groups. **C**, Risk score distribution of 11 lncRNAs (picture above) and survival status of patients (picture below). **D**, Heat map of 11 lncRNAs expression profiles.



**Figure 3.** ROC curves of 11 lncRNAs associated with LUSC prognosis. **A-K**, ROC curves of ADAMTS9-AS2, LINC00628, LINC00961, RP11-374A4.1, CTB-147C22.9, RP11-697M17.2, RP11-76C10.2, RP6-24A23.7, RP11-279O17.1, RP11-133K1.11, RP11-1085N6.3.



**Figure 4.** Kaplan-Meier estimates OS of LUSC patients based on individual lncRNA. **A-K**, Kaplan-Meier survival curves of ADAMTS9-AS2, LINC00628, LINC00961, RP11-374A4.1, CTB-147C22.9, RP11-697M17.2, RP11-76C10.2, RP6-24A23.7, RP11-279O17.1, RP11-133K1.11, RP11-1085N6.3.

shorter OS, indicating that these 7 lncRNAs were high-risk genes. However, patients with high expression of the other 4 lncRNAs, including RP6-24A23.7 (Figure 4H), RP11-279O17.1 (Figure 4I), RP11-133K1.11 (Figure 4J) and RP11-1085N6.3 (Figure 4K) tended to have longer OS, indicating that these four lncRNAs were low-risk genes. The results were consistent with the prediction of multivariate Cox regression analysis.

## Discussion

LUSC is a common type of lung cancer that causes approximately 400,000 deaths worldwide each year<sup>19</sup>. In recent years, as the number of LUSC patients with poor prognosis has increased, it is of vital importance to select more effective prognostic biomarkers to predict LUSC patients' survival<sup>20</sup>.

MiRNAs and lncRNAs can be used as biomarkers for LUSC diagnosis and prognosis<sup>21-23</sup>. In recent years, people have begun to study the role of lncRNAs in cancer diagnosis and prognosis, and found that lncRNAs are involved in carcinogenic or tumor suppressive pathways<sup>24</sup>. The expression of lncRNAs is related to the occurrence, metastasis and prognosis of tumors<sup>25</sup>. Many scholars<sup>26</sup> have tended to study individual lncRNA biomarker, but other have shown that combination of several lncRNAs have better predictive ability than individual lncRNA. Since the expression of lncRNA is relatively lower, it may easily lead to errors when using individual lncRNA as a biomarker<sup>27</sup>. Therefore, in this study, we conducted a comprehensive analysis of the lncRNAs expressed in LUSC and identified a total of 2085 DElncRNAs. 342 lncRNAs were selected by univariate Cox regression analysis. The LASSO regression model identified 11 lncRNAs associated with the prognosis of LUSC, namely ADAMTS9-AS2, LINC00628, LINC00961, RP11-374A4.1, CTB-147C22.9, RP11-697M17.2, RP11-76C10.2, RP6-24A23.7, RP11-279O17.1, RP11-133K1.11 and RP11-1085N6.3.

Then, we proved that the risk scoring model constructed by 11 lncRNAs had good performance. The prognostic ability of the model was independent with patients' gender, age, smoking status and pathological stage, and could distinguish survival status preferably, which was completely different from the previous lncRNA-based prognosis models of LUSC<sup>6,17,18</sup>. The reasons may be as follows. Firstly, we used LASSO penalty regression model to analyze the 342 genes before

multivariate Cox regression analysis. Unlike traditional stepwise regression, LASSO algorithm can simultaneously analyze all independent variables and tends to select the most influential variables<sup>28</sup>. A penalty term was added into regularization method. Then, the penalty function was used for analysis after regularization and the coefficient of the variables with less impact will become zero<sup>29</sup>. So, this approach is more accurate when dealing with large datasets<sup>30</sup>. Secondly, we also conducted independent survival analysis for each lncRNA. The AUC of individual lncRNA was smaller than that of the 11-lncRNA model, which further verified the superiority of the model. Among these 11 lncRNAs, ADAMTS9-AS2, LINC00628, LINC00961, RP11-374A4.1, CTB-147C22.9, RP11-697M17.2, RP11-76C10.2 were high-risk genes, while RP6-24A23.7, RP11-279O17.1, RP11-133K1.11, RP11-1085N6.3 were low-risk genes. Previous investigations have shown that ADAMTS9-AS2 is associated with the prognosis of gastric cancer<sup>31</sup>, salivary adenoid cystadenocarcinoma<sup>32</sup>, and glioma<sup>33</sup>, and the expression of LINC00628 is associated with the prognosis of breast cancer<sup>34</sup>. LINC00961 has also been considered as a tumor suppressor gene and can be used as a prognostic biomarker for NSCLC<sup>35</sup> and renal cell carcinoma<sup>36</sup>.

Literature searching on PubMed showed that it was the first report exhibiting the correlation between RP11-374A4.1, CTB-147C22.9, RP11-697M17.2, RP11-76C10.2, RP6-24A23.7, RP11-279O17.1, RP11-133K1.11, RP11-1085N6.3 and the prognosis of LUSC. Our study provided a 11-lncRNA risk scoring model with high accuracy, which indicated that these 11 lncRNAs can be used as prognostic biomarkers for LUSC patients. The study provides new insight as well as a more effective way for LUSC prognosis. More studies will be conducted to evaluate the biological functions of these lncRNAs by studying their effects on cell proliferation and apoptosis, further advancing the development of tumor prognosis.

## Conclusions

We constructed a new 11-lncRNA model to predict the prognosis of LUSC. The model had good performance and could accurately distinguish the prognosis risk of patients with high and low risk, thus providing a potential tool for the prognosis of LUSC.

### Conflict of Interest

The Authors declare that they have no conflict of interests.

### References

- 1) OSER MG, NIEDERST MJ, SEQUIST LV, ENGELMAN JA. Transformation from non-small-cell lung cancer to small-cell lung cancer: molecular drivers and cells of origin. *Lancet Oncol* 2015; 16: e165-172.
- 2) DRILON A, REKHTMAN N, LADANYI M, PAIK P. Squamous-cell carcinomas of the lung: emerging biology, controversies, and the promise of targeted therapy. *Lancet Oncol* 2012; 13: e418-426.
- 3) BECK JT, ISMAIL A, TOLOMEO C. Targeting the phosphatidylinositol 3-kinase (PI3K)/AKT/mammalian target of rapamycin (mTOR) pathway: an emerging treatment strategy for squamous cell lung carcinoma. *Cancer Treat Rev* 2014; 40: 980-989.
- 4) MILLER KD, SIEGEL RL, LIN CC, MARIOTTO AB, KRAMER JL, ROWLAND JH, STEIN KD, ALTERI R, JEMAL A. Cancer treatment and survivorship statistics, 2016. *CA Cancer J Clin* 2016; 66: 271-289.
- 5) ANTAL CE, HUDSON AM, KANG E, ZANCA C, WIRTH C, STEPHENSON NL, TROTTER EW, GALLEGOS LL, MILLER CJ, FURNARI FB, HUNTER T, BROGNARD J, NEWTON AC. Cancer-associated protein kinase C mutations reveal kinase's role as tumor suppressor. *Cell* 2015; 160: 489-502.
- 6) LUO D, DENG B, WENG M, LUO Z, NIE X. A prognostic 4-lncRNA expression signature for lung squamous cell carcinoma. *Artif Cells Nanomed Biotechnol* 2018; 46: 1207-1214.
- 7) LI J, LI Z, ZHENG W, LI X, WANG Z, CUI Y, JIANG X. LncRNA-ATB: an indispensable cancer-related long noncoding RNA. *Cell Prolif* 2017; 50.
- 8) ULITSKY I, BARTEL DP. lincRNAs: genomics, evolution, and mechanisms. *Cell* 2013; 154: 26-46.
- 9) TERASHIMA M, ISHIMURA A, WANNA-UDOM S, SUZUKI T. MEG8 long noncoding RNA contributes to epigenetic progression of the epithelial-mesenchymal transition of lung and pancreatic cancer cells. *J Biol Chem* 2018; 293: 18016-18030.
- 10) SUN Q, HAO Q, PRASANTH KV. Nuclear long non-coding RNAs: Key regulators of gene expression. *trends genet* 2018; 34: 142-157.
- 11) VAN HEESCH S, VAN ITERSON M, JACOBI J, BOYMANS S, ESSERS PB, DE BRUIJN E, HAO W, MACINNES AW, CUPPEN E, SIMONIS M. Extensive localization of long non-coding RNAs to the cytosol and mono- and polyribosomal complexes. *Genome Biol* 2014; 15: R6.
- 12) RASHID F, SHAH A, SHAN G. Long non-coding RNAs in the cytoplasm. *Genomics Proteomics Bioinformatics* 2016; 14: 73-80.
- 13) XIE Y, ZHANG Y, DU L, JIANG X, YAN S, DUAN W, LI J, ZHAN Y, WANG L, ZHANG S, LI S, WANG L, XU S, WANG C. Circulating long noncoding RNA act as potential novel biomarkers for diagnosis and prognosis of non-small cell lung cancer. *Mol Oncol* 2018; 12: 648-658.
- 14) PENG W, WANG J, SHAN B, PENG Z, DONG Y, SHI W, HE D, CHENG Y, ZHAO W, ZHANG C, LI B, DUAN C. Diagnostic and prognostic potential of circulating long non-coding rnas in non small cell lung cancer. *Cell Physiol Biochem* 2018; 49: 816-827.
- 15) OSIENSKA MA, JAGODZINSKI PP. Long non-coding RNA as potential biomarkers in non-small-cell lung cancer: what do we know so far? *Biomed Pharmacother* 2018; 101: 322-333.
- 16) RICCIUTI B, MENCARONI C, PAGLIALUNGA L, PACIULLO F, CRINO L, CHIARI R, METRO G. Long noncoding RNAs: new insights into non-small cell lung cancer biology, diagnosis and therapy. *Med Oncol* 2016; 33: 18.
- 17) LUO Y, XUAN Z, ZHU X, ZHAN P, WANG Z. Long non-coding RNAs RP5-821D11.7, APCDD1L-AS1 and RP11-277P12.9 were associated with the prognosis of lung squamous cell carcinoma. *Mol Med Rep* 2018; 17: 7238-7248.
- 18) TANG RX, CHEN WJ, HE RQ, ZENG JH, LIANG L, LI SK, MA J, LUO DZ, CHEN G. Identification of a RNA-Seq based prognostic signature with five lncRNAs for lung squamous cell carcinoma. *Oncotarget* 2017; 8: 50761-50773.
- 19) CANCER GENOME ATLAS RESEARCH NETWORK. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 2012; 489: 519-525.
- 20) LI R, YANG YE, JIN J, ZHANG MY, LIU X, LIU XX, YIN YH, QU YO. Identification of lncRNA biomarkers in lung squamous cell carcinoma using comprehensive analysis of lncRNA mediated ceRNA network. *Artif Cells Nanomed Biotechnol* 2019; 47: 3246-3258.
- 21) XU B, CHEN Q, YUE C, LAN L, JIANG J, SHEN Y, LU B. Prognostic value of IL-6R mRNA in lung adenocarcinoma and squamous cell carcinoma. *Oncol Lett* 2018; 16: 2935-2948.
- 22) ISIK Z, ERCAN ME. Integration of RNA-Seq and RP-PA data for survival time prediction in cancer patients. *Comput Biol Med* 2017; 89: 397-404.
- 23) QI L, GAO C, FENG F, ZHANG T, YAO Y, WANG X, LIU C, LI J, LI J, SUN C. MicroRNAs associated with lung squamous cell carcinoma: new prognostic biomarkers and therapeutic targets. *J Cell Biochem* 2019; 120: 18956-18966.
- 24) ZHOU M, GUO M, HE D, WANG X, CUI Y, YANG H, HAO D, SUN J. A potential signature of eight long non-coding RNAs predicts survival in patients with non-small cell lung cancer. *J Transl Med* 2015; 13: 231.
- 25) SPIZZO R, ALMEIDA MI, COLOMBATTI A, CALIN GA. Long non-coding RNAs and cancer: a new frontier of translational research? *Oncogene* 2012; 31: 4577-4587.
- 26) HUANG GW, XUE YJ, WU ZY, XU XE, WU JY, CAO HH, ZHU Y, HE JZ, LI CO, LI EM, XU LY. A three-lncRNA signature predicts overall survival and disease-free survival in patients with esophageal

- squamous cell carcinoma. *BMC Cancer* 2018; 18: 147.
- 27) WU B, WANG K, FEI J, BAO Y, WANG X, SONG Z, CHEN F, GAO J, ZHONG Z. Novel three lncRNA signature predicts survival in patients with pancreatic cancer. *Oncol Rep* 2018; 40: 3427-3437.
- 28) FRIEDMAN J, HASTIE T, TIBSHIRANI R. Regularization Paths for generalized linear models via coordinate descent. *J Stat Softw* 2010; 33: 1-22.
- 29) GAO J, KWAN PW, SHI D. Sparse kernel learning with LASSO and Bayesian inference algorithm. *Neural Netw* 2010; 23: 257-264.
- 30) MCNEISH DM. Using lasso for predictor selection and to assuage overfitting: a method long overlooked in behavioral sciences. *Multivariate Behav Res* 2015; 50: 471-484.
- 31) CAO B, LIU C, YANG G. Down-regulation of lncRNA ADAMTS9-AS2 contributes to gastric cancer development via activation of PI3K/Akt pathway. *Biomed Pharmacother* 2018; 107: 185-193.
- 32) XIE S, YU X, LI Y, MA H, FAN S, CHEN W, PAN G, WANG W, ZHANG H, LI J, LIN Z. Upregulation of lncRNA ADAMTS9-AS2 Promotes salivary adenoid cystic carcinoma metastasis via PI3K/Akt and MEK/Erk signaling. *Mol Ther* 2018; 26: 2766-2778.
- 33) YAO J, ZHOU B, ZHANG J, GENG P, LIU K, ZHU Y, ZHU W. A new tumor suppressor lncRNA ADAMTS9-AS2 is regulated by DNMT1 and inhibits migration of glioma cells. *Tumour Biol* 2014; 35: 7935-7944.
- 34) CHEN DQ, ZHENG XD, CAO Y, HE XD, NIAN WQ, ZENG XH, LIU XY. Long non-coding RNA LINC00628 suppresses the growth and metastasis and promotes cell apoptosis in breast cancer. *Eur Rev Med Pharmacol Sci* 2017; 21: 275-283.
- 35) HUANG Z, LEI W, TAN J, HU HB. Long noncoding RNA LINC00961 inhibits cell proliferation and induces cell apoptosis in human non-small cell lung cancer. *J Cell Biochem* 2018; 119: 9072-9080.
- 36) CHEN D, ZHU M, SU H, CHEN J, XU X, CAO C. LINC00961 restrains cancer progression via modulating epithelial-mesenchymal transition in renal cell carcinoma. *J Cell Physiol* 2019; 234: 7257-7265.