

Identification of a progression-associated long non-coding RNA signature for predicting the prognosis of lung squamous cell carcinoma

YANYAN WANG, FANG YANG and YONGZHI ZHUANG

Department of Medical Oncology, Daqing Oilfield General Hospital, Daqing, Heilongjiang 163001, P.R. China

Received August 2, 2016; Accepted April 7, 2017

DOI: 10.3892/etm.2017.5571

Abstract. Long non-coding RNAs (lncRNAs) have been indicated to have prognostic roles in various cancer types. However, the association between lncRNAs and lung squamous cell carcinoma (LSCC) progression, and the prognostic value of lncRNAs as a marker for early detection of LSCC have not been systematically investigated. The present study performed a genome-wide comparative analysis in order to determine the expression profiles of 10,207 lncRNAs to investigate the expression patterns between patients with early stages of LSCC (stage I-II) and those with late-stage disease (stage III-IV). It was found that 114 lncRNAs were significantly differentially expressed between late- and early-stage LSCC and therefore associated with the progression of the malignancy. By focusing on progression-associated lncRNAs, eight lncRNAs were found to be significantly associated with overall survival of LSCC patients according to univariate Cox proportional hazards regression analysis. These eight prognostic lncRNAs were integrated into a progression-associated eight-lncRNA signature, which stratified patients into two groups with significantly different overall survival (median survival, 2.25 vs. 7.68 years; $P=2.06 \times 10^{-5}$). In addition, the prognostic value of the progression-associated eight-lncRNA signature was independent of known clinical factors. Functional analysis suggested that eight prognostic lncRNAs may be involved in adipocytokine signaling pathway and glycerophospholipid metabolism. Taken together, the progression-associated eight-lncRNA signature identified by our study not only represents a candidate prognostic biomarker for LSCC patients but also provides insight into the molecular mechanisms in the progression of LSCC.

Introduction

Lung cancer, characterized by uncontrolled cell growth in lung tissues, remains the leading cause of cancer-associated mortality worldwide. In China, lung cancer is one of the most common cancer types. In 2010, an estimate of 605,900 patients were newly diagnosed with lung cancer in China, accounting for 19.59% of all new cancer cases, and 486,600 deaths were due to lung cancer (1). Lung cancer is classified into two major types based on the cells' histological characteristics: Non-small cell lung cancer (NSCLC) and SCLC. NSCLC is the most common sub-type of lung cancer. NSCLC can be classified into three main histological subtypes: Lung adenocarcinoma (LUAD), squamous-cell carcinoma (LSCC) and large-cell carcinoma (LLCC). Nearly 30% of lung cancer cases are LSCC (2). The 5-year survival rate varied greatly, ranging from ~60% in patients with stage I to <5% for patients with stage IV (3). Therefore, there is an urgent need to identify progression-associated biomarkers for early detection and develop a novel and improved predictive tools for LSCC prognosis.

It is well known that only a small fraction of the human genome (<2%) contains protein-coding genes, while >98% of the genome may be transcribed to non-coding RNAs (ncRNAs) that lack the capacity to encode proteins (4). ncRNAs may be divided into two main types based on their size: Shorter ncRNAs, such as micro (mi)RNAs and long non-coding RNAs (lncRNAs). lncRNAs, the largest class of ncRNAs, are mRNA-like transcripts of >200 nucleotides in length, which lack protein-coding capacity (5). A large number of studies have suggested that lncRNAs have important roles in nearly all aspects of gene regulation at the transcriptional, post-transcriptional and epigenetic level (6,7). Increasing evidence has revealed a close association between lncRNAs and human cancers (8,9). Dysregulated lncRNA expression has been implicated in various human cancers (8-12). For instance, a well-characterized lncRNA, metastasis-associated lung adenocarcinoma transcript 1, has been found to be frequently upregulated in numerous cancer types and was indicated to have clinical value as a therapeutic target or prognostic marker in cancer (12,13). Another lncRNA, HOTAIR, is overexpressed in lung and ovarian cancer and its levels are correlated with metastasis and poor prognosis (14,15). Studies have suggested a clinical application

Correspondence to: Dr Yanyan Wang, Department of Medical Oncology, Daqing Oilfield General Hospital, 9 Zhongkang Street, Saertu, Daqing, Heilongjiang 163001, P.R. China
E-mail: ywangedu@163.com

Key words: lung squamous cell carcinoma, long non-coding RNAs, prognosis, expression profiles

of lncRNAs as prognostic markers in some cancer types, including oesophageal squamous cell carcinoma, NSCLC, breast cancer, LUAD, B-cell lymphoma, ovarian cancer, glioblastoma multiforme, pancreatic cancer, multiple myeloma and colorectal cancer (16-28). However, the prognostic value of lncRNAs as a marker for the early detection of LSCC has not been systematically investigated.

The present study aimed to identify progression-associated lncRNAs by a genome-wide analysis of lncRNA expression profiles between patients with early stages of LSCC (stage I-II) and those with late-stage disease (stage III-IV). A novel progression-associated eight-lncRNA signature was developed and validated to improve the prognosis prediction for patients with LSCC.

Materials and methods

Patient dataset. Clinical and pathological data of patients with LSCC were retrieved from the Cancer Genome Atlas (TCGA) data portal (<https://tcga-data.nci.nih.gov/tcga/>). A total of 109 LSCC patients with lncRNA expression data, including 30 females and 79 males, and 87 stage-I/II and 21 stage-III/IV patients, were enrolled in the study. Detailed clinical information on LSCC patients enrolled in this study is listed in Table I.

Acquisition and analysis of lncRNA expression profiles of LSCC patients. The lncRNA expression dataset was derived from the study of Du *et al* (29), repurposing the probes from an Affymetrix Human Exon 1.0 ST microarray. In brief, the probes of the Affymetrix Human 1.0 ST array were uniquely mapped to lncRNA using the latest annotations of lncRNA with a computational pipeline. A total of 10,207 lncRNA-encoding genes with at least 4 probes were obtained for further analysis. The expression value of lncRNA was obtained by summarizing the background-corrected intensity of all probes corresponding to this lncRNA and was standardized using the quantile-normalized method and an empirical Bayes method.

Expression profiles of lncRNAs between LSCC patients at early stages and those with late-stage disease were compared and the differentially expressed lncRNAs were identified using the significance analysis of microarrays method. Those lncRNAs with a P-value of <0.01 and a fold change of >1.5 or <0.67 were considered as differentially expressed lncRNAs. Unsupervised hierarchical clustering of LSCC patients and lncRNAs was performed with the R platform (version 3.2.5; <https://www.r-project.org>) using the euclidean distance and complete linkage method.

Construction of progression-based lncRNA prognostic signature. To identify lncRNAs associated with the outcome for LSCC patients, univariate Cox regression analysis was performed to examine the association between expression levels of progression-associated lncRNAs and patients' overall survival. Those lncRNAs with a P-value of <0.05 were selected as prognostic lncRNAs whose expression levels were significantly associated with patients' overall survival. A progression-based lncRNA prognostic signature was then constructed by linear combination of the expression levels of

Table I. Clinical and pathological characteristics of patients with lung squamous cell carcinoma in the Cancer Genome Atlas cohort (n=109).

Covariate	Value
Age, (years)	66.747 (39.301-84.816)
Stage	
I, II	87 (79.8)
III, IV	21 (19.3)
NA	1 (0.9)
Gender	
Male	79 (72.5)
Female	30 (27.5)
Smoking	
Yes	92 (84.4)
No	17 (15.6)
Vital status	
Alive	64 (58.7)
Dead	45 (41.3)
Values are expressed as n (%) or mean (range).	

eight prognostic lncRNAs with the multivariate Cox regression coefficient as the weight.

Statistical analysis. Kaplan-Meier survival curves were used to assess the difference in overall survival between the high-risk and the low-risk group, and statistical significance was evaluated using the two-sided log-rank test. Multivariate Cox regression analysis and stratification analysis were performed to determine whether the lncRNA signature was independent of other clinical features. Time-dependent receiver operating characteristic (ROC) curve analysis was also performed to evaluate the sensitivity and specificity of the lncRNA signature for survival prediction at five years. The area under the curve (AUC) value was calculated from the ROC curve. All analyses were performed using the R platform (version 3.2.5) and Bio-conductor (<https://www.bioconductor.org/>).

Functional enrichment analysis. Functional enrichment analysis was performed to predict the functions of eight prognostic lncRNAs at the Gene ontology (GO) and Kyoto encyclopedia of genes and genomes (KEGG) pathway levels using the Database for Annotation, Visualization and Integrated Discovery Bioinformatics Resources 6.8 (Beta; <https://david-d.ncicrf.gov/>) (30). The results of the enrichment analysis were obtained limited to GO terms in the 'Biological Process' (GOTERM-BP-FAT) and KEGG pathway categories using the functional annotation clustering and functional annotation chart options with the human whole genome as the background. Enriched GO functional annotation clusters and KEGG pathways with P<0.05 and enrichment score >1.5 were considered as potential functions of prognostic lncRNAs.

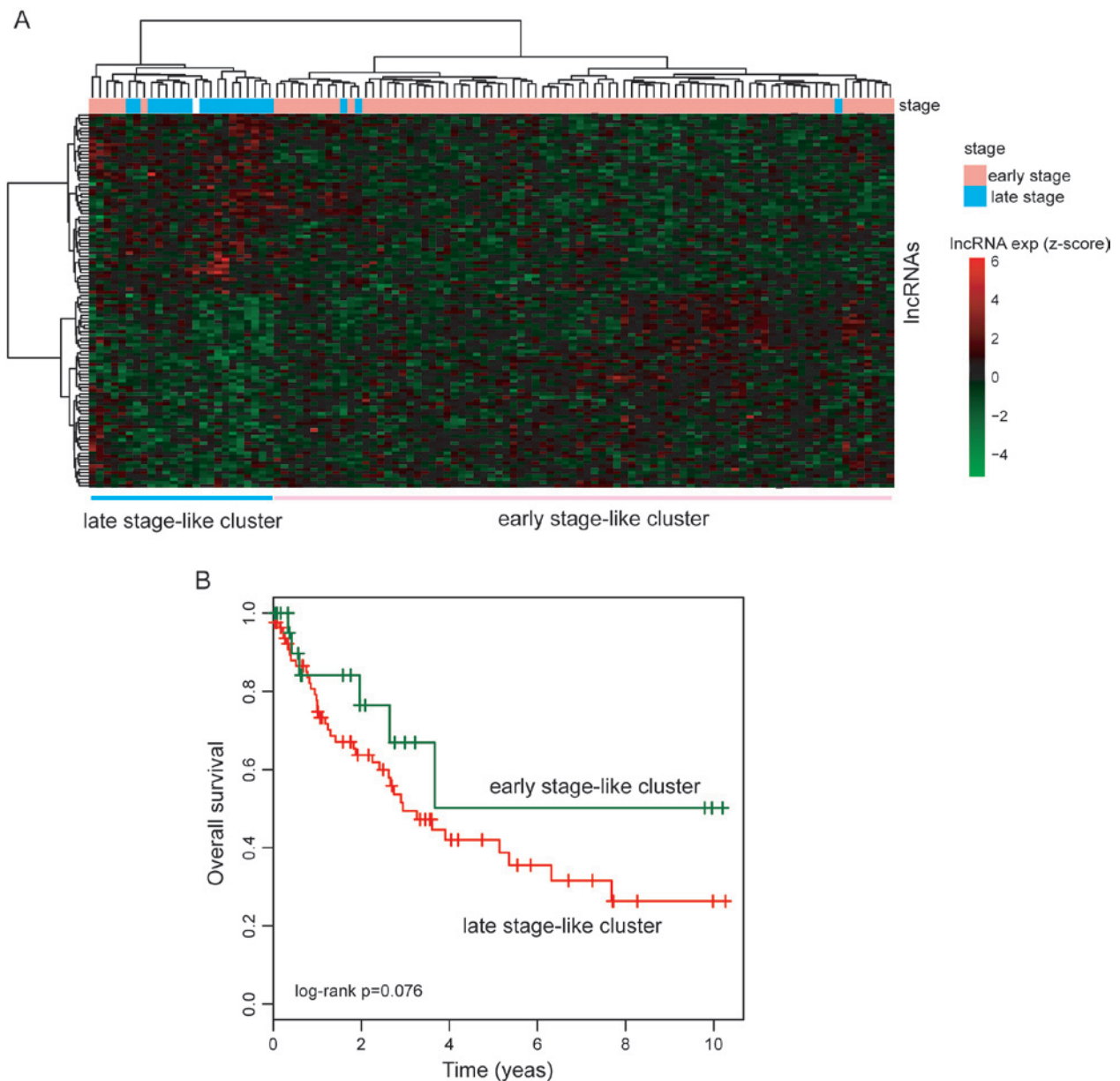


Figure 1. Association between lncRNAs and progression and prognosis of LSCC. (A) Hierarchical cluster analysis of LSCC patients based on the expression profiles of 114 differentially expressed lncRNAs. (B) Kaplan-Meier survival curves of patients with LSCC in the early stage-like cluster and late stage-like cluster. LSCC, lung squamous cell carcinoma; lncRNA, long non-coding RNA.

Results

Identification of progression-associated lncRNAs in LSCC.

In order to identify LSCC progression-associated lncRNAs, a significance analysis of microarrays was first performed to compare expression profiles of patients with early stages of LSCC (stage I-II) with those with late-stage disease (stage III-IV). Comparative analysis of lncRNA expression profiles between patients with early-stage LSCC (stage I-II) and those with late-stage disease (stage III-IV) revealed obviously different expression patterns and identified a total of 114 differentially expressed lncRNAs ($P < 0.01$ and fold change > 1.5 or < 0.67).

Unsupervised hierarchical clustering of 109 LSCC patients according to the expression patterns of these 114 differentially expressed lncRNAs showed two distant patient clusters, which were highly correlated with patients' stage

status ($P = 6.132 \times 10^{-14}$, Fisher's exact test; Fig. 1A). The early stage-like cluster contained 98.8% of patients with early stages of LSCC and the late stage-like cluster contained 75% of patients with late-stage disease, as shown in Fig. 1A. Furthermore, a survival analysis revealed a marginally significant difference in overall survival between patients with early-stage LSCC and those with late-stage disease ($P = 0.076$, log-rank test). The median survival time of patients with late-stage disease was 2.95 years, which was significantly lower than that of patients with early-stage LSCC (Fig. 1B).

Identification of prognostic lncRNAs from progression-associated lncRNAs. To identify potential progression-associated prognostic lncRNAs, the expression data of 114 differentially expressed lncRNAs between patients with early- and those with late-stage LSCC were subjected

Table II. Progression-associated lncRNAs significantly associated with overall survival of patients with lung squamous cell carcinoma in the Cancer Genome Atlas cohort.

lncRNA ID	Ensembl name	Genomic location	P-value ^a	Hazard ratio (95% CI) ^a	Coefficient of univariate analysis ^a
ENSG00000249297	RP11-97E7.2	Chr 4: 171,602,683-171,638,763 (-)	0.022	0.669 (0.474-0.944)	-0.402
ENSG00000230314	ELOVL2-AS1	Chr 6: 11,043,524-11,078,226 (+)	0.025	0.727 (0.551-0.960)	-0.318
ENSG00000234171	RNASEH1-AS1	Chr 2: 3,558,492-3,561,745 (+)	0.049	0.739 (0.547-0.999)	-0.303
ENSG00000226337	RP11-274B18.4	Chr 9: 68,541,036-68,644,442 (+)	0.046	0.731 (0.537-0.994)	-0.313
ENSG00000204055	RP11-247A12.2	Chr 9: 129,176,771-129,210,548 (+)	0.047	0.758 (0.577-0.996)	-0.277
ENSG00000234182	RP11-118K6.2	Chr 10: 3,010,531-3,013,111 (+)	0.019	0.679 (0.491-0.938)	-0.387
ENSG00000259628	RP11-467H10.2	Chr 7: 77,043,721-77,198,626 (+)	0.034	1.407 (1.026-1.930)	0.341
ENSG00000236896	RP11-535C21.3	Chr 9: 97,986,551-97,987,656 (-)	0.026	0.689 (0.497-0.956)	-0.372

^aDerived from univariate Cox proportional hazards regression analysis. Chr, chromosome; lncRNA, long non-coding RNA; CI, confidence intervals.

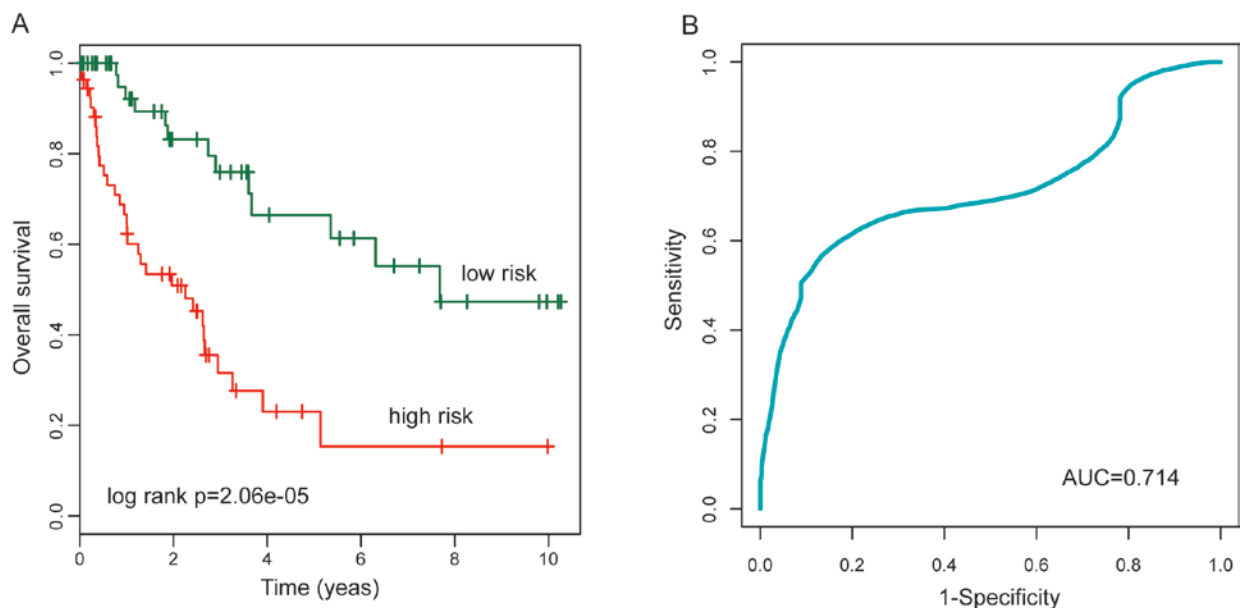


Figure 2. Validation of the progression-associated eight-lncRNA signature in the Cancer Genome Atlas cohort. (A) Kaplan-Meier survival curves of patients with lung squamous cell carcinoma in the high-risk and the low-risk group. (B) Time-dependent receiver operating characteristics curves of the progression-associated eight-lncRNA signature at five years. AUC, area under curve; lncRNA, long non-coding RNA.

to univariate Cox proportional hazards regression analysis with overall survival as the dependent variable. A set of eight lncRNAs that was significantly correlated with patients' overall survival was identified ($P < 0.05$; Table II). Of these, seven lncRNAs had negative coefficients, indicating that their high expression was associated with longer survival. The remaining one lncRNAs (RP11-467H10.2) had a positive coefficient, indicating that its high expression was associated with shorter survival.

Construction and validation of progression-associated eight-lncRNA signature in the TCGA cohort. A progression-associated eight-lncRNA signature was developed by linear combination of the expression levels of the eight prognostic lncRNAs weighted by the multivariate Cox regression coefficient as follows: Risk score = $(-0.2410 \times \text{expression value of RP11-97E7.2}) + (-0.3645 \times \text{expression value of ELOVL2-AS1}) + (-0.2909 \times \text{expression value of RNASEH1-AS1}) + (-0.2296 \times \text{expression value of RP11-274B18.4}) + (-0.3476 \times \text{expression value of RP11-247A12.2}) + (-0.3684 \times \text{expression value of RP11-118K6.2}) + (0.1972 \times \text{expression value of RP11-467H10.2}) + (-0.1977 \times \text{expression value of RP11-535C21.3})$. This allowed for assignment of a risk score to each patient according to the progression-associated eight-lncRNA signature. Using the median risk score as the cutoff (0.1066), the LSCC patients were stratified into the high-risk group ($n=55$) and the low-risk group ($n=54$). Kaplan-Meier curves for the high-risk and low-risk groups are shown in Fig. 2A. Patients in the high-risk group had a significantly shorter median overall survival than those in the low-risk group (median survival, 2.25 vs. 7.68 years; $P=2.06 \times 10^{-5}$,

value of RP11-97E7.2) + $(-0.3645 \times \text{expression value of ELOVL2-AS1}) + (-0.2909 \times \text{expression value of RNASEH1-AS1}) + (-0.2296 \times \text{expression value of RP11-274B18.4}) + (-0.3476 \times \text{expression value of RP11-247A12.2}) + (-0.3684 \times \text{expression value of RP11-118K6.2}) + (0.1972 \times \text{expression value of RP11-467H10.2}) + (-0.1977 \times \text{expression value of RP11-535C21.3})$. This allowed for assignment of a risk score to each patient according to the progression-associated eight-lncRNA signature. Using the median risk score as the cutoff (0.1066), the LSCC patients were stratified into the high-risk group ($n=55$) and the low-risk group ($n=54$). Kaplan-Meier curves for the high-risk and low-risk groups are shown in Fig. 2A. Patients in the high-risk group had a significantly shorter median overall survival than those in the low-risk group (median survival, 2.25 vs. 7.68 years; $P=2.06 \times 10^{-5}$,

Table III. Uni- and multivariate Cox regression analysis of the progression-associated eight-lncRNA signature and survival in the Cancer Genome Atlas cohort.

Variable	Univariate analysis			Multivariate analysis		
	HR	95% CI	P-value	HR	95% CI	P-value
Eight-lncRNA signature (high vs. low)	3.847	1.985-7.454	6.57×10^{-5}	3.778	1.903-7.498	1.45×10^{-4}
Age (>65 vs. ≤65 years)	1.035	1.000-1.071	0.049	1.043	1.005-1.0816	0.027
Stage (III/IV vs. I/II)	1.115	0.496-2.507	0.793	1.351	0.591-3.092	0.476
Gender (male vs. female)	1.733	0.857-3.507	0.126	1.471	0.711-3.043	0.299
Smoking status (yes vs. no)	1.223	0.436-3.434	0.702	1.059	0.365-3.071	0.917

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

log-rank test; Fig. 2A). Univariate analysis revealed a significant association between the progression-associated eight-lncRNA signature and overall survival, in which the hazard ratio (HR) for overall survival with comparison between the high-vs. the low-risk group was 3.847 [95% confidence interval (CI)=1.985-7.454; $P=6.57 \times 10^{-5}$; Table III). The five-year time-dependent ROC curve analysis for the progression-associated eight-lncRNA signature determined an AUC of 0.714 at five years (Fig. 2B). These results demonstrated that this progression-associated eight-lncRNA signature allows for an improved prognosis prediction of patients with LSCC.

Fig. 3A presents a risk score distribution of LSCC patients calculated using the progression-associated eight-lncRNA signature ranked according to the risk scores. It was found that in the high-risk group, more mortalities occurred than in the low-risk group (Fig. 3B). Furthermore, seven prognostic lncRNAs tended to be linked with a good prognosis, as their high expression was associated with longer survival, and only one lncRNA was a risk-associated lncRNA whose high expression was associated with shorter survival (Fig. 3C).

Independent prognostic value of the progression-associated eight-lncRNA signature. To determine whether the prognostic value of the progression-associated eight-lncRNA signature is independent of other clinical variables, a multivariate Cox regression analysis was performed in each patient cohort, which included risk scores, age, stage, gender and smoking status as co-variables. It was found that the progression-associated eight-lncRNA signature (HR=3.778, 95% CI=1.903-7.498; $P=1.45 \times 10^{-4}$) and age (HR=1.043, 95% CI=1.005-1.082; $P=0.027$) were significantly and independently associated with overall survival in the multivariate analysis (Table III). A further stratification analysis for patient age was performed to test whether the progression-associated eight-lncRNA signature is independent of age. All LSCC patients included in the TCGA dataset were stratified by age into either a younger stratum (age, ≤65 years) or an older stratum (age, >65 years). Within the younger stratum, the progression-associated eight-lncRNA signature was able to classify patients into the high-risk and the low-risk group with significantly different overall survival ($P=0.045$, log-rank test; Fig. 4A). Similarly, older patients were divided into a high-risk group (n=38) and a low-risk group (n=28) according to the threshold of the median risk score. Patients in the high-risk group

had a significantly shorter median survival time than those in the low-risk group (median survival, 1.3 vs. 7.68 years; $P=8.35 \times 10^{-5}$, log-rank test; Fig. 4B). The results of the multivariable Cox regression analysis and stratification analysis thus demonstrated that the prognostic value of the progression-associated eight-lncRNA signature is independent of other clinical factors linked with the overall survival of patients with LSCC.

Functional analysis of the progression-associated eight-lncRNA signature. To explore the potential functional roles of the lncRNAs in the progression-associated eight-lncRNA signature in the progression of LSCC, lncRNA function was predicted through GO and KEGG pathway functional enrichment analysis for mRNAs co-expressed with lncRNAs. The co-expression associations between eight prognostic lncRNAs and mRNAs were first determined by calculating the Pearson correlation coefficient of paired lncRNA and mRNA expression profiles and 1,246 mRNAs positively correlated (ranked top 0.5%) with at least one of eight prognostic lncRNAs were identified. GO analysis suggested that the co-expressed mRNAs were most significantly enriched in two GO functional annotation clusters (mainly involved in regulation of transcription and RNA processing). Functional enrichment analysis of the KEGG pathway revealed that the co-expressed mRNAs were most significantly enriched in two KEGG pathways, namely the adipocytokine signaling pathway and glycerophospholipid metabolism. Functional analysis suggested that lncRNAs in the progression-associated eight-lncRNA signature may affect critical biological pathways and processes involved in LSCC progression.

Discussion

Although recent advances in the understanding of the molecular mechanisms of LSCC have greatly improved the prognosis prediction for patients diagnosed with LSCC, existing molecular biomarkers focus on protein-coding genes or miRNAs (31,32). lncRNAs are a class of ncRNAs, which are expressed in a more tissue-, cell type- and cancer type-specific manner than protein-coding genes, implicating their intrinsic advantages over protein-coding RNAs in diagnosis and prognosis (33). A recent study investigated the expression of lncRNAs in LSCC in a cohort of 16 patients and identified

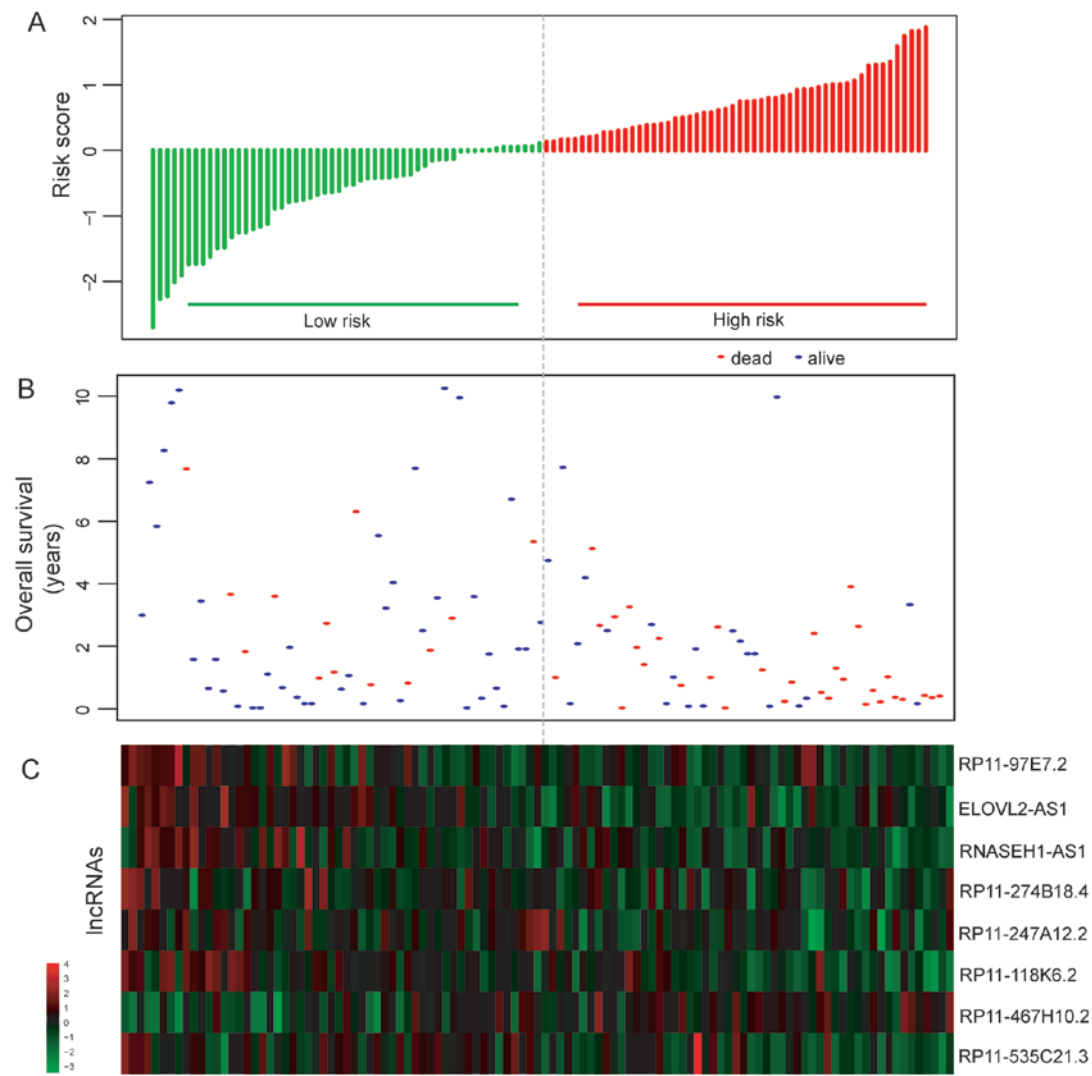


Figure 3. LncRNA risk score analysis in the Cancer Genome Atlas cohort. (A) LncRNA risk score distribution; (B) patients' overall survival status and time; (C) heatmap of the lncRNA expression profiles. lncRNA, long non-coding RNA.

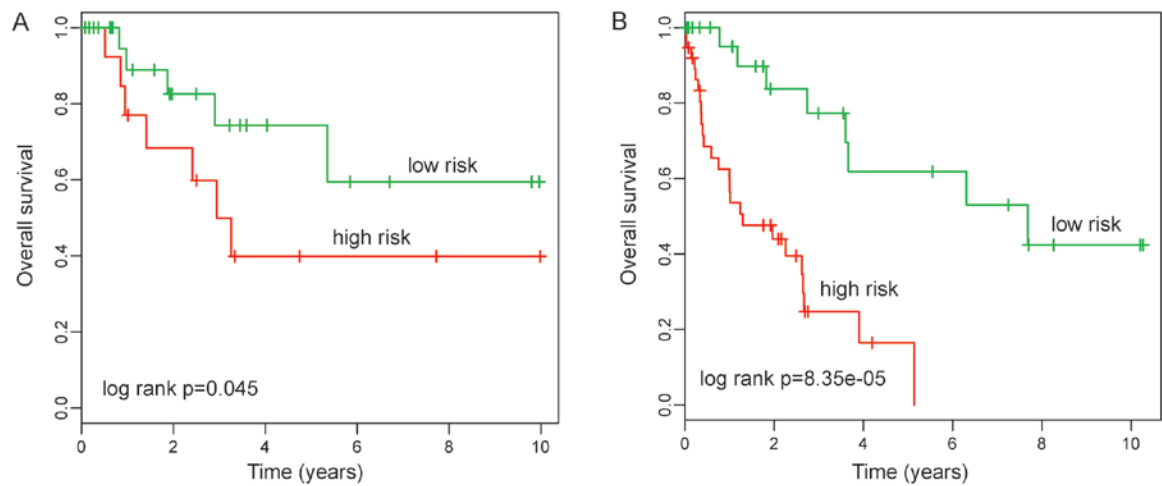


Figure 4. Survival prediction in younger and older patients. (A) Kaplan-Meier survival curves of younger patients (age, ≤65 years). (B) Kaplan-Meier survival curves of older patients (age, >65 years).

significantly differentially expressed lncRNAs in LSCC compared to matched normal lung tissues (34). However, the

prognostic value of lncRNAs as a marker for early detection of LSCC has not been systematically investigated.

In the present study, a genome-wide comparative analysis of expression profiles of 10,207 lncRNA-encoding genes between patients with early-stage LSCC (stage I-II) and those with late-stage disease (stage III-IV) was performed. Through the above analysis of tumor samples in different stages, a total of 114 lncRNAs were identified to be significantly and differentially expressed in patients with late-vs. early-stage LSCC. Among the differentially expressed lncRNAs, 55 were upregulated and 59 were downregulated in patients with late-stage disease (stage III-IV), indicating that these differentially expressed lncRNAs may be associated with the progression of LSCC. With a focus on progression-associated lncRNAs, univariate Cox proportional hazards regression analysis was used to identify eight prognostic lncRNAs significantly associated with the overall survival of LSCC patients. To facilitate their clinical application, these eight lncRNAs were integrated into a progression-associated eight-lncRNA signature based on the linear combination of expression data of eight prognostic lncRNAs and weighted by the regression coefficients from the multivariate Cox regression analysis. To confirm the survival prediction power of the progression-based lncRNA prognostic signature, the median risk score was used as the threshold to stratify LSCC patients into the high-risk and the low-risk group. Patients in the low-risk group tended to have prolonged overall survival, whereas patients in the high-risk group tended to have shortened survival, suggesting that the progression-based lncRNA prognostic signature is a powerful tool for prognostic prediction of patients with LSCC. Furthermore, the progression-based lncRNA prognostic signature is independent of other clinical factors, including age, stage, gender and smoking status.

Although the number of known lncRNAs has grown over the past years, relatively few of lncRNAs have been well studied. For instance, only 184 human lncRNAs whose function is known are included in the lncRNAdb v2.0 database (35). It has been suggested that lncRNAs function by interacting with their partner molecules. Therefore, the function of lncRNAs may be predicted using the functions of their co-expressed mRNAs (36). To gain functional insight into eight prognostic lncRNAs, the present study performed a functional enrichment analysis for co-expressed mRNAs to infer potential biological processes and pathways of lncRNAs. It was found that these lncRNAs were mainly involved in the regulation of transcription, RNA processing, the adipocytokine signaling pathway and glycerophospholipid metabolism. Numerous studies have highlighted the important roles of the adipocytokine signaling pathway in the development of cancer (37,38). A metabolomic profiling analysis by Hoang *et al* (39) found that the glycerophospholipid metabolism was altered in NSCLC. The functional analysis of the present study based on the co-expressed mRNAs with lncRNAs suggested that eight prognostic lncRNAs may be involved in critical biological pathways and processes associated with tumor progression.

Taken together, the progression-associated eight-lncRNA signature identified in the present study not only represents a candidate biomarker for the prognosis of LSCC patients but also provided insights into the molecular mechanisms associated with the progression of LSCC.

References

- Chen W, Zheng R, Zeng H and Zhang S: Epidemiology of lung cancer in China. *Thorac Cancer* 6: 209-215, 2015.
- Verlicchi A, Karachaliou N, Lazzari C, Servat CC, Chaib I, Marin S, Gkoutakos A, Pilotto S, Serrano JL and Rosell R: 72P differential expression profile of lung squamous cell carcinoma (LSCC) cell lines as a mean to predict drug interaction effects. *J Thorac Oncol* 11 (4 Suppl): S85-S86, 2016.
- Tanoue LT and Detterbeck FC: New TNM classification for non-small-cell lung cancer. *Expert Rev Anticancer Ther* 9: 413-423, 2009.
- ENCODE Project Consortium, Birney E, Stamatoyannopoulos JA, Dutta A, Guigó R, Gingeras TR, Margulies EH, Weng Z, Snyder M, Dermitzakis ET, *et al*: Identification and analysis of functional elements in 1% of the human genome by the ENCODE pilot project. *Nature* 447: 799-816, 2007.
- Ponting CP, Oliver PL and Reik W: Evolution and functions of long noncoding RNAs. *Cell* 136: 629-641, 2009.
- Cao J: The functional role of long non-coding RNAs and epigenetics. *Biol Proced Online* 16: 11, 2014.
- Kornienko AE, Guenzl PM, Barlow DP and Pauler FM: Gene regulation by the act of long non-coding RNA transcription. *BMC Biol* 11: 59, 2013.
- Sun J, Shi H, Wang Z, Zhang C, Liu L, Wang L, He W, Hao D, Liu S and Zhou M: Inferring novel lncRNA-disease associations based on a random walk model of a lncRNA functional similarity network. *Mol Biosyst* 10: 2074-2081, 2014.
- Zhou M, Wang X, Li J, Hao D, Wang Z, Shi H, Han L, Zhou H and Sun J: Prioritizing candidate disease-related long non-coding RNAs by walking on the heterogeneous lncRNA and disease network. *Mol Biosyst* 11: 760-769, 2015.
- Chen G, Wang Z, Wang D, Qiu C, Liu M, Chen X, Zhang Q, Yan G and Cui Q: lncRNADisease: A database for long-non-coding RNA-associated diseases. *Nucleic Acids Res* 41 (Database Issue): D983-D986, 2013.
- Cheng L, Shi H, Wang Z, Hu Y, Yang H, Zhou C, Sun J and Zhou M: IntNetLncSim: An integrative network analysis method to infer human lncRNA functional similarity. *Oncotarget* 7: 47864-47874, 2016.
- Tian X and Xu G: Clinical value of lncRNA MALAT1 as a prognostic marker in human cancer: Systematic review and meta-analysis. *BMJ Open* 5: e008653, 2015.
- Ren D, Li H, Li R, Sun J, Guo P, Han H, Yang Y and Li J: Novel insight into MALAT-1 in cancer: Therapeutic targets and clinical applications (Review). *Oncol Lett* 11: 1621-1630, 2016.
- Wang J, Chen D, He X, Zhang Y, Shi F, Wu D, Chen J, Zhang Y, Zhao F and Dou J: Downregulated lincRNA HOTAIR expression in ovarian cancer stem cells decreases its tumorigenesis and metastasis by inhibiting epithelial-mesenchymal transition. *Cancer Cell Int* 15: 24, 2015.
- Loewen G, Jayawickramarajah J, Zhuo Y and Shan B: Functions of lncRNA HOTAIR in lung cancer. *J Hematol Oncol* 7: 90, 2014.
- Hu Y, Chen HY, Yu CY, Xu J, Wang JL, Qian J, Zhang X and Fang JY: A long non-coding RNA signature to improve prognosis prediction of colorectal cancer. *Oncotarget* 5: 2230-2242, 2014.
- Zhou M, Zhao H, Xu W, Bao S, Cheng L and 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* 16: 16, 2017.
- Zhou M, Zhang Z, Zhao H, Bao S, Cheng L and Sun J: An immune-related six-lncRNA signature to improve prognosis prediction of glioblastoma multiforme. *Mol Neurobiol*: May 19, 2017 (Epub ahead of print).
- Zhou M, Zhong L, Xu W, Sun Y, Zhang Z, Zhao H, Yang L and Sun J: Discovery of potential prognostic long non-coding RNA biomarkers for predicting the risk of tumor recurrence of breast cancer patients. *Sci Rep* 6: 31038, 2016.
- Zhou M, Diao Z, Yue X, Chen Y, Zhao H, Cheng L and Sun J: Construction and analysis of dysregulated lncRNA-associated ceRNA network identified novel lncRNA biomarkers for early diagnosis of human pancreatic cancer. *Oncotarget* 7: 56383-56394, 2016.
- Zhou M, Wang X, Shi H, Cheng L, Wang Z, Zhao H, Yang L and Sun J: Characterization of long non-coding RNA-associated ceRNA network to reveal potential prognostic lncRNA biomarkers in human ovarian cancer. *Oncotarget* 7: 12598-12611, 2016.

22. Zhou M, Sun Y, Xu W, Zhang Z, Zhao H, Zhong Z and Sun J: Comprehensive analysis of lncRNA expression profiles reveals a novel lncRNA signature to discriminate nonequivalent outcomes in patients with ovarian cancer. *Oncotarget* 7: 32433-32448, 2016.
23. Sun J, Cheng L, Shi H, Zhang Z, Zhao H, Wang Z and Zhou M: A potential panel of six-long non-coding RNA signature to improve survival prediction of diffuse large-B-cell lymphoma. *Sci Rep* 6: 27842, 2016.
24. Zhou M, Xu W, Yue X, Zhao H, Wang Z, Shi H, Cheng L and Sun J: Relapse-related long non-coding RNA signature to improve prognosis prediction of lung adenocarcinoma. *Oncotarget* 7: 29720-29738, 2016.
25. Zhou M, Zhao H, Wang Z, Cheng L, Yang L, Shi H, Yang H and Sun J: Identification and validation of potential prognostic lncRNA biomarkers for predicting survival in patients with multiple myeloma. *J Exp Clin Cancer Res* 34: 102, 2015.
26. Sun J, Chen X, Wang Z, Guo M, Shi H, Wang X, Cheng L and Zhou M: A potential prognostic long non-coding RNA signature to predict metastasis-free survival of breast cancer patients. *Sci Rep* 5: 16553, 2015.
27. Zhou M, Guo M, He D, Wang X, Cui Y, Yang H, Hao D and Sun J: A potential signature of eight long non-coding RNAs predicts survival in patients with non-small cell lung cancer. *J Transl Med* 13: 231, 2015.
28. Li J, Chen Z, Tian L, Zhou C, He MY, Gao Y, Wang S, Zhou F, Shi S, Feng X, *et al*: lncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. *Gut* 63: 1700-1710, 2014.
29. Du Z, Fei T, Verhaak RG, Su Z, Zhang Y, Brown M, Chen Y and Liu XS: Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. *Nat Struct Mol Biol* 20: 908-913, 2013.
30. Huang da W, Sherman BT and Lempicki RA: Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* 4: 44-57, 2009.
31. Zhu CQ, Strumpf D, Li CY, Li Q, Liu N, Der S, Shepherd FA, Tsao MS and Jurisica I: Prognostic gene expression signature for squamous cell carcinoma of lung. *Clin Cancer Res* 16: 5038-5047, 2010.
32. Fukumoto I, Hanazawa T, Kinoshita T, Kikkawa N, Koshizuka K, Goto Y, Nishikawa R, Chiyomaru T, Enokida H, Nakagawa M, *et al*: MicroRNA expression signature of oral squamous cell carcinoma: Functional role of microRNA-26a/b in the modulation of novel cancer pathways. *Br J Cancer* 112: 891-900, 2015.
33. Cheetham SW, Gruhl F, Mattick JS and Dinger ME: Long noncoding RNAs and the genetics of cancer. *Br J Cancer* 108: 2419-2425, 2013.
34. Wang Y, Qian CY, Li XP, Zhang Y, He H, Wang J, Chen J, Cui JJ, Liu R, Zhou H, *et al*: Genome-scale long noncoding RNA expression pattern in squamous cell lung cancer. *Sci Rep* 5: 11671, 2015.
35. Quek XC, Thomson DW, Maag JL, Bartonicek N, Signal B, Clark MB, Gloss BS and Dinger ME: lncRNADB v2.0: Expanding the reference database for functional long noncoding RNAs. *Nucleic Acids Res* 43 (Database Issue): D168-D173, 2015.
36. Liao Q, Liu C, Yuan X, Kang S, Miao R, Xiao H, Zhao G, Luo H, Bu D, Zhao H, *et al*: Large-scale prediction of long non-coding RNA functions in a coding-non-coding gene co-expression network. *Nucleic Acids Res* 39: 3864-3878, 2011.
37. Masaki T and Yoshimatsu H: Obesity, adipocytokines and cancer. *Transl Oncogenomics* 3: 45-52, 2008.
38. Vansaun MN: Molecular pathways: Adiponectin and leptin signaling in cancer. *Clin Cancer Res* 19: 1926-1932, 2013.
39. Hoang CD, Wu M, Xu Y, Merrit R, Whyte R, Shrager J and Peltz G: Altered glycerophospholipid metabolism in non-small cell lung carcinoma. *Cancer Res* 70: 5574-5574, 2010.



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