

## Identification of long noncoding RNAs for the detection of early stage lung squamous cell carcinoma by microarray analysis

Zule Cheng<sup>1,2</sup>, Yanan Bai<sup>1,2</sup>, Ping Wang<sup>1</sup>, Zhenhua Wu<sup>1,2</sup>, Lin Zhou<sup>1,2</sup>, Ming Zhong<sup>3</sup>, Qinghui Jin<sup>1</sup>, Jianlong Zhao<sup>1</sup>, Hailei Mao<sup>3</sup>, Hongju Mao<sup>1</sup>

<sup>1</sup>State Key Laboratory of Transducer Technology, Shanghai Institute of Microsystem and Information Technology, Chinese Academy of Science, Shanghai 200050, China

<sup>2</sup>University of Chinese Academy of Sciences, Beijing 100039, China

<sup>3</sup>Departments of Anesthesiology and Critical Care Medicine, Zhongshan Hospital, Fudan University, Shanghai 200032, China

**Correspondence to:** Jianlong Zhao, **email:** jlzhao@mail.sim.ac.cn  
Hailei Mao, **email:** mao.hailei@zs-hospital.sh.cn  
Hongju Mao, **email:** hjmao@mail.sim.ac.cn

**Keywords:** long noncoding RNA, biomarker, diagnosis, early stage lung squamous cell carcinoma, microarray

**Received:** October 20, 2016

**Accepted:** December 27, 2016

**Published:** January 05, 2017

### ABSTRACT

The aberrant expressions of long noncoding RNAs have been reported in numerous cancers, which have facilitated the cancer diagnosis. However, the expression profile of lncRNAs in early stage lung squamous cell carcinoma has not been well discussed. The present study aimed to examine the expression profile of lncRNAs in early stage lung squamous cell carcinoma and identify lncRNA biomarkers for diagnosis. Through high-throughput lncRNA microarray, we screened thousands of aberrantly expressed lncRNAs and mRNAs in early stage lung squamous cell carcinoma tissues compared to their corresponding adjacent nontumorous tissues. Bioinformatics analyses were used to investigate the functions of aberrantly expressed mRNAs and their associated lncRNAs. After that, in order to identify lncRNA biomarkers for early detection, candidate lncRNA biomarkers were selected based on our established filtering pipeline and validated by real-time quantitative polymerase chain reaction on a total of 63 pairs of tumor samples. Five lncRNAs were finally identified which were able to distinguish early stage tumor and normal samples with high sensitivity (92%) and specificity (83%). These results imply that lncRNAs may be powerful biomarker for early diagnosis.

### INTRODUCTION

Non-small cell lung cancer (NSCLC) is a heterogeneous group of tumors with three subtypes: lung adenocarcinoma (ADC), lung squamous cell carcinoma (LSCC), and large-cell carcinoma. The majority of NSCLC patients is at an advanced disease stage or present with distant metastasis at the time of diagnosis, leaving them with limited treatment options. This delay in diagnosis results in an overall 5-year survival rate of less than 16% [1]. Diagnosing NSCLC earlier would likely reduce overall mortality and increase the rate of remission. As one of the NSCLC subtypes, LSCC accounts for nearly 30% of NSCLC worldwide [2]. Thus, there is considerable interest in developing early detection for LSCC. Biomarkers such as squamous cell

carcinoma antigen (SCC-Ag) and CYFRA 21-1 are classic biomarkers commonly used in the diagnoses of LSCCs [3]. However, the current lacking of sufficient diagnostic sensitivity and specificity have limited their usefulness in the early detection of LSCC.

Long noncoding RNAs (lncRNAs) are transcripts longer than 200 nucleotides with little or no protein-coding potential [4]. Dysregulation of lncRNAs is associated with a number of cancer-related processes [5], including epigenetic regulation [6], microRNA silencing [7], DNA damage and cell cycle control [8, 9]. Therefore, lncRNAs have the potential as powerful biomarkers for detecting carcinomas. Some well-studied lncRNAs have been detected in NSCLCs, such as CARLo-5 [10], CCAT2 [11], SOX2OT [12], MEG3 [13] and TUG1 [14]. CARLo-5, CCAT2 and SOX2OT are functioned

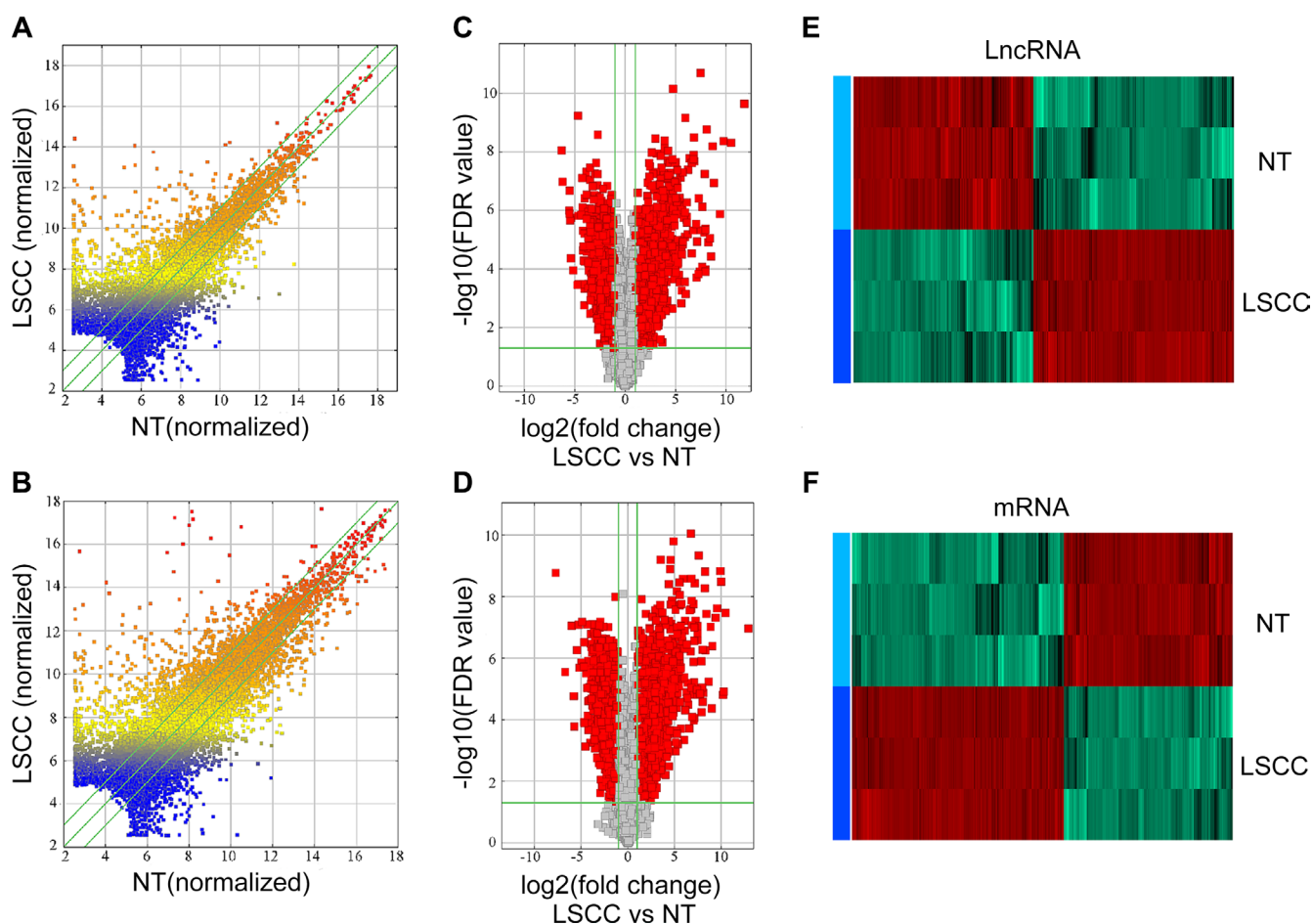
as onco-lncRNAs and overexpressed in NSCLCs. SOX2OT is a negative prognostic indicator for NSCLC, and the expression level of SOX2OT was found higher in LSCC than ADC [12]. Meanwhile, MEG3 and TUG1 are tumor-suppressor lncRNAs that are downregulated in NSCLCs. With the development of high-throughput technology, increasing number of novel diagnostic and prognostic lncRNA expression signatures are identified, particularly in lung ADC [15, 16]. However, little is known about the lncRNA expression profile during early stage LSCC.

In this work, we used microarray analysis to examine the lncRNA expression profile in early stage LSCC samples against corresponding nontumorous tissue (NT) samples. The signatures of differentially expressed lncRNAs and mRNAs were analyzed. Especially, a novel panel of five lncRNAs was identified which could distinguish early stage LSCC sample from NT samples.

## RESULTS

### Landscape of expressed lncRNAs and mRNAs in LSCC

In total, 30,586 lncRNAs and 26,109 mRNAs were detected in three pairs of early stage LSCC samples by microarray analysis. Differences of lncRNA and mRNA expression in LSCC samples relative to NT controls are shown in Figure 1A and 1B. A total of 2,668 lncRNAs and 3,728 mRNAs had significantly aberrant expression in all three LSCC samples, as evaluated by volcano plot filtering (Figure 1C and 1D). Of the 2,668 aberrantly expressed lncRNAs in the LSCC samples, 1,407 were upregulated and 1,261 were downregulated. Among the 3,728 differentially expressed mRNAs (DE-mRNA), 2,074 were upregulated and 1,654 were downregulated. Hierarchical clustering analysis (Figure 1E and 1F) resulted in distinct expression signatures of both lncRNAs and mRNAs in the LSCC and NT samples.



**Figure 1: Expression profiles of lncRNAs and mRNAs in LSCCs compared with NT samples.** (A and B) Scatter-plot of variation in expression of lncRNAs and mRNAs between early stage LSCC tissues and NT samples. X and Y axes are the mean normalized signal values ( $\log_2$  scaled). The green lines are fold change lines (the default fold change value given is 2.0). (C and D) The Volcano plot of differentially expressed lncRNAs and mRNAs in LSCC relative to NT samples. The vertical green lines represent 2.0-fold changes up and down and the horizontal green line represents a FDR-value of 0.05. The red points in the plot represent the differentially expressed lncRNAs with statistical significance, X axes are the Fold change values ( $\log_2$  scaled), Y axes are the FDR-values ( $\log_2$  scaled). (E) and (F) Hierarchical clustering analysis of lncRNAs and mRNAs.

## Co-expression analysis of DE-lncRNAs and associated coding genes in LSCC

Several studies have reported that some lncRNAs can activate or inhibit their associated coding genes by transcriptional interference or chromatin modification [17] [18]. Of the 2,668 DE-lncRNAs, 1,086 genic DE-lncRNAs were localized with their protein coding genes and 286 of these lncRNAs' associated coding genes were on the list of DE-mRNAs (Figure 2A). Among these 286 coding genes, 83% expressed in the same direction with their corresponding associated lncRNAs, while the rest read in the opposite direction. Furthermore, from the microarray data, we identified 113 enhancer-like lncRNAs and 171 intergenic lncRNAs that were co-aberrantly expressed with their nearby target genes (Figure 2B). We found that some of these lncRNAs have multiple associated target genes. This feature may relate to their locations on the chromosomes.

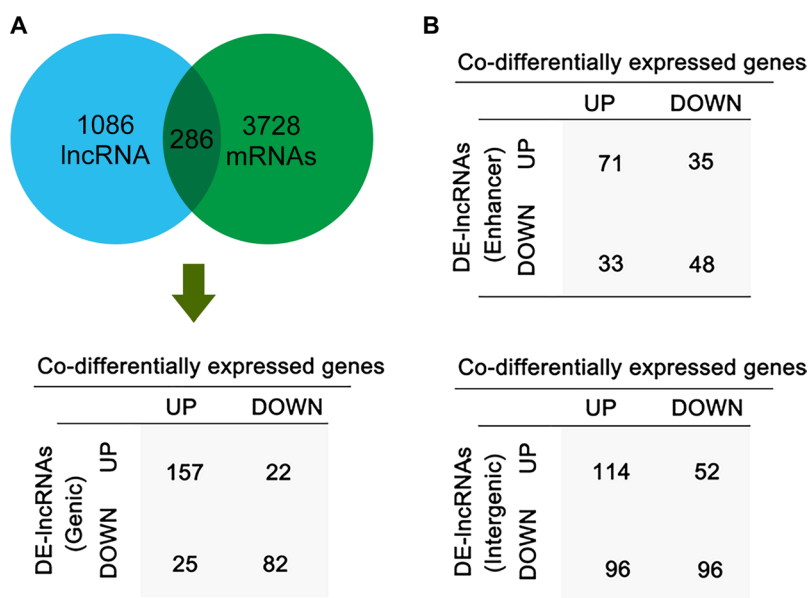
### Gene ontology and pathway analysis

To systematically assign putative functions to protein coding genes, we performed GO and KEGG pathway analyses. GO analysis revealed that the upregulated genes were involved in GO terms like nuclear lumen, mitotic cell cycle processes, and protein binding (Figure 3A, right). The downregulated genes were involved in responses to stimuli, extracellular regions, and IgG binding (Figure 3A, left). KEGG pathway analysis was performed by mapping 3,728 DE-mRNAs to KEGG pathways. It was found that upregulated DE-mRNAs were fell disproportionately into several cancer-related pathways (Figure 3B, right). The top three enriched pathways were involved in the cell

cycle, p53 signaling and pathways in cancer. ErbB and Wnt signaling pathways also appeared in the top twenty enriched pathways. All these pathways described are closely associated with NSCLCs [19, 20]. By contrast, downregulated mRNAs were disproportionately fell into pathways like *Staphylococcus aureus* infection, phagosomes, and lysosomes, which may be responsible for various infection (Figure 3B, left). Next, we performed the GO and Pathway analysis to the group of DE-lncRNA associated DE-mRNAs which was generated from former analysis. We found these DE-mRNAs were enriched in pathways like Cell cycle, Transcriptional misregulation in cancer and Choline metabolism in cancer (Supplementary Tables 1 and 2). We speculated that some DE-lncRNAs might be either directly or indirectly involved in these pathways.

### Evaluate the diagnostic potential of lncRNAs in early stage LSCC

In order to evaluate the diagnostic potential of lncRNAs in early stage LSCC, we adopt an established method to select candidate biomarkers from 2,668 DE-lncRNAs [16]. 16 upregulated lncRNAs and 10 downregulated lncRNAs were selected (Supplementary Table 3). Quantitative polymerase chain reaction (RT-qPCR) was used to further validate the expression levels of the 26 candidate lncRNAs in a training set of 24 pairs of early stage LSCC and NT samples (Table 1). Among these 26 candidates, five were validated which had significantly different expressions in the early stage LSCC samples (Table 2, Figure 4A–4E), while the other 21 candidates had no significant differences (Supplementary Table 4). These five lncRNAs showed the same trends in



**Figure 2: The expression relationship of DE-lncRNAs and their associated coding gene in LSCC. (A)** The analysis of genic lncRNAs and their associated coding genes. The table listed the expression correlation of genic lncRNAs and their associated coding gene. **(B)** The expression correlation of enhance-like lncRNAs (left) and intergenic lncRNA (right) with their nearby coding genes.

**Table 1: Clinical pathologic characteristics of the training set of 24 LSCC patients and the independent validation cohort of 39 LSCC patients**

	Training set (24 pairs)	Validation set (39 pairs)
<b>Age</b>		
Mean ± SD	59.8 ± 5.6	61.5 ± 6.9
<b>Gender</b>		
Female	2	2
Male	22	37
<b>TNM stage</b>		
I	11	13
II	13	10
III	0	16
<b>Tumor grade (differentiation)</b>		
Well	5	4
Moderate	17	26
Poor	2	9

**Table 2: Expression levels of the five lncRNAs and their AUC values in training set and validation set**

Group	LncRNA	Mean (SD) in NT	Mean (SD) in LSCC	p-Value	AUC (SE)	Sensitivity (%)	Specificity (%)
<b>Training Set:</b>	ENST00000453324	1.19 (1.63)	3.74 (3.94)	0.001481	0.76 (0.07)	69	84
	NR_028500	0.98 (1.00)	2.36 (2.17)	0.001293	0.70 (0.07)	73	65
	UC011CLY.2	3.37 (3.36)	9.08 (5.85)	0.000013	0.77 (0.06)	84	66
	NR_046326	2.16 (2.44)	6.91 (5.40)	0.000040	0.79 (0.06)	84	74
	ENST00000441841	4.78 (5.20)	13.48 (11.01)	0.000288	0.76 (0.06)	73	73
<b>Validation Set:</b>	ENST00000453324	2.02 (2.18)	5.42 (4.49)	0.000010	0.77 (0.05)	64	80
	NR_028500	2.32 (2.66)	6.41 (6.16)	0.000191	0.73 (0.06)	75	68
	UC011CLY.2	3.94 (4.02)	12.64 (10.02)	0.000003	0.80 (0.05)	73	80
	NR_046326	4.12 (4.53)	10.65 (8.36)	0.000026	0.77 (0.05)	79	71
	ENST00000441841	8.10 (10.80)	20.10 (16.92)	0.000006	0.80 (0.07)	86	70

Abbreviations: AUC: area under the receiver operating characteristic curve; SD: standard deviation; SE: standard error. Both the p value and AUC were obtained using the  $\beta$ -actin normalized values.

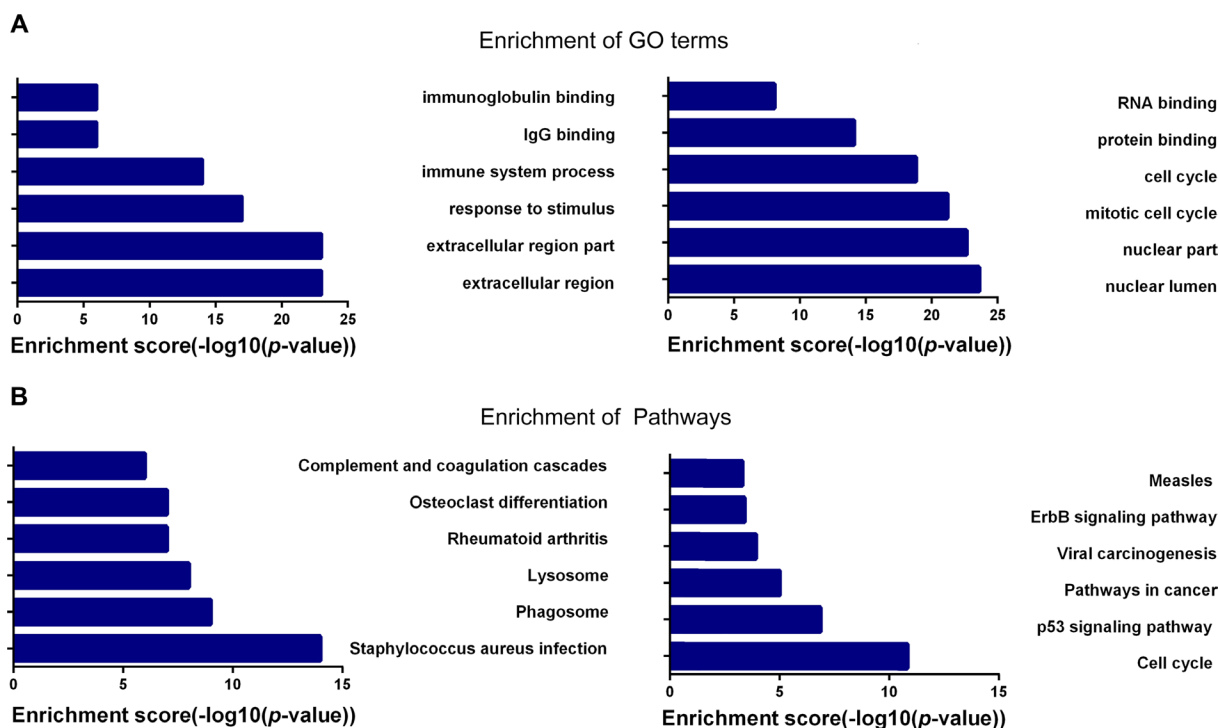
regulatory direction as the microarray data (Figure 4F), which supports a strong consistency between the qPCR and the microarray data. Then, ROC curve analysis was used to evaluate the distinguish capability of these five lncRNAs (Table 2). Combination of these five lncRNAs produced a higher AUC value (0.905, Figure 5A) with corresponding increases in sensitivity (92%) and specificity (83%).

We then used another independent set of 39 pairs of LSCC and NT samples (Table 1) to validate these five lncRNAs. Results yielded a similar result as training set (Table 2). Combination of five lncRNAs showed an AUC value of 0.912 (Figure 5B) with 85% sensitivity and 87% specificity. We also examined the expression levels of these five lncRNAs in advanced stage LSCC samples,

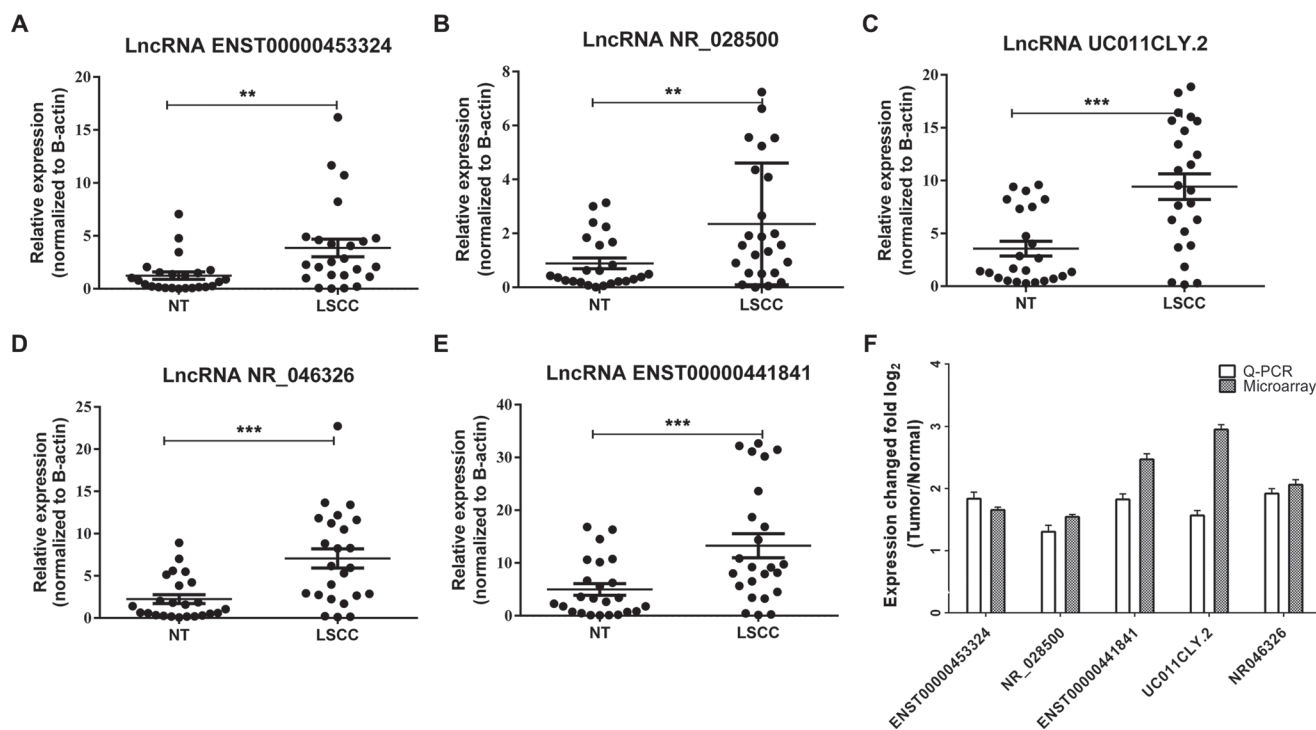
the results showed that the differences were significant between NT and advanced stage LSCC samples, but were not significant between early and advanced stage LSCC samples (Supplementary Tables 5 and 6).

## DISCUSSION

People diagnosed with clinical stage I and II LSCC have a 60% to 40% 5-year rate of survival, however those with clinical stage III or IV disease have a less than 5% 5-year rate of survival [21]. Thus, early diagnosis may help to improve the 5-year rate of survival. A thorough understanding of the molecular aberrations of LSCC would assist in the biomarkers identification and early detection. Using a high-throughput lncRNA array, we



**Figure 3: Enrichment analysis of pathways and GO terms for differentially expressed mRNAs.** The top 6 gene ontology analysis (A) and top 6 pathways (B) that exhibited significant differences between LSCC and NT samples are listed (left and right panel show the coding genes of downregulated and upregulated, respectively).

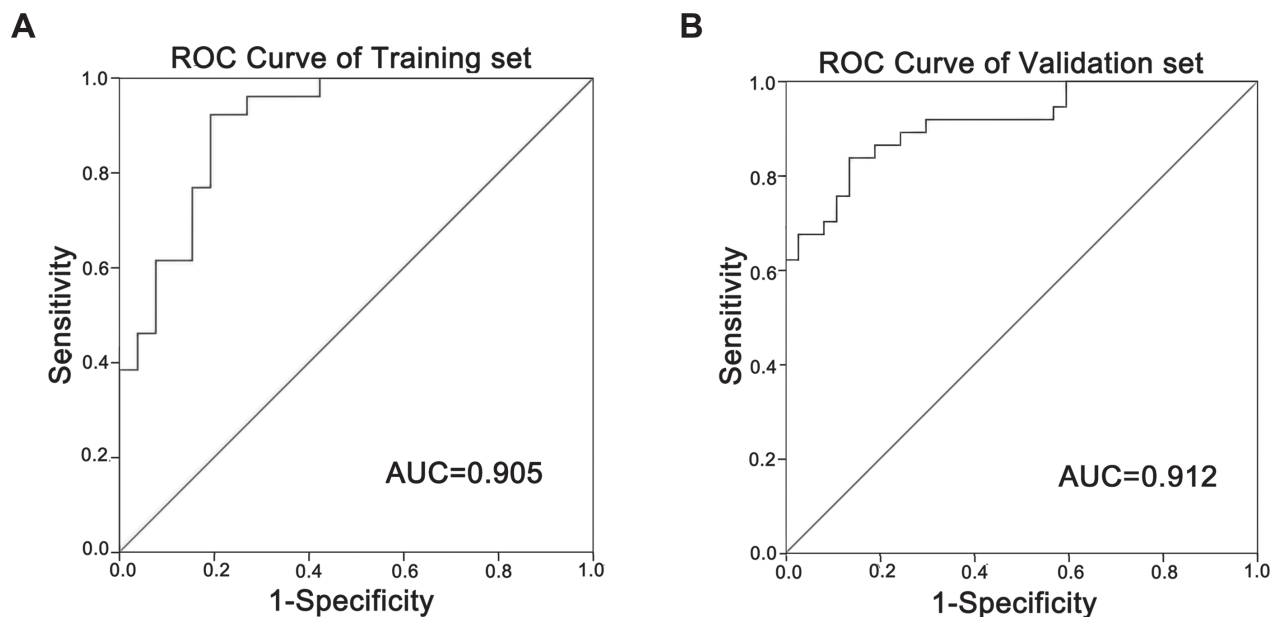


**Figure 4: RT-qPCR validation reveals significant lncRNA expression differences in LSCC samples compare to the adjacent normal samples.** (A–E) The expression level of five selected lncRNAs in 24 LSCC samples against corresponding adjacent NT samples. All the differences of five lncRNAs between LSCC samples and NT samples were significant. (F) Comparison of the expression change fold of five candidate lncRNAs in three paired patients' microarray results and 24 paired patients' qPCR results. The heights of the columns in the chart represent the mean expression value of log<sub>2</sub> fold changes (tumor/normal) for each of the five validated lncRNAs in the microarray and RT-qPCR data; the bars represent standard errors. \*\* statistically significant at:  $p < 0.01$  (two-tailed  $t$ -test); \*\*\* statistically significant at:  $p < 0.001$  (two-tailed  $t$ -test).

identified thousands of DE-lncRNAs and DE-mRNAs in three pairs of early stage LSCC and corresponding adjacent NT samples. The potential functions of DE-lncRNAs were then discussed through function analysis and co-expression analysis, which were mainly based on the involvement of lncRNAs' functions with their associated protein-coding genes [18, 22–25]. Results indicated that part of these DE-lncRNAs might be involved in some cancer-related pathway, but further studies were still needed to validate these speculations.

Increasing evidences have suggested that cancer-associated lncRNAs have the potential as diagnostic and prognostic biomarkers [26–28]. We then evaluated the diagnostic ability of lncRNAs in early stage LSCC. Using a training and validation set in a separate independent cohort, a panel of five lncRNAs was identified and optimized. Combination of these five lncRNAs shows a considerable power in distinguishing early stage LSCC samples from NT samples. The functions of these five lncRNAs remain unknown, however their associated coding genes may be involved in some cancer-related processes. For example, intron sense overlapping lncRNA ENST00000453324, also named as LINC000969, is located on the sense strand of chromosome 3 and is transcribed from the intron area of gene mucin 20 (MUC20). Overexpression of MUC20 can enhance EGF-induced phosphorylation of EGFR and STAT3 [29]. However, we didn't observe the aberrant expression of MUC20 in our microarray data, indicating the function of LINC000969 may not related to gene MUC20. NR\_028500 is located on the sense strand of chromosome 11 and is a transcript variant of gene lactate dehydrogenase A (LDHA). Interestingly, NR\_028500 associated gene LDHA is also found overexpressed in LSCC tissues

(fold change > 2,  $P < 0.001$ ). Protein encoded by this gene catalyzes the conversion of L-lactate and NAD to pyruvate and NADH in the final step of anaerobic glycolysis. Importantly, LDHA is upregulated in tumors and promotes cellular proliferation and tumorigenesis [30–32]. The co-expression of NR\_028500 and LDHA implies that NR\_028500 may be involved in the anaerobic glycolysis of tumor cells. ENST00000441841 and NR\_046326 are two enhancer-like lncRNAs, which may increase transcription of their nearby coding genes [33]. NR\_046326, also named as PCAT6 (prostate cancer associated transcript 6), is located on the sense strand of chromosome 1, and is first reported highly expressed in prostate cancer [34]. Study revealed that PCAT6 might enhance the expression of its neighbor gene KLHL12 and activate the Wnt- $\beta$  catenin pathway indirectly [33]. Recent studies identified that PCAT6 was significantly upregulated in lung adenocarcinoma tissues [35], and was positively associated with metastasis and cellular proliferation in lung cancer [36]. In this work, overexpression of PCAT6 was also identified in early stage LSCC tissues, implying that PCAT6 was closely related to the tumorigenesis in LSCC. LncRNA ENST00000441841 is located on the sense strand of chromosome X. Gene PLAC1 (placenta-specific protein 1 precursor) is on the upstream of it and defined as a neighbor coding gene (distance < 300 kb). PLAC1 is expressed exclusively in trophoblast cells in the mammalian placenta, and is involved in trophoblast invasion and migration [37]. Recent studies have detected high PLAC1 expression levels in a numbers of human solid tumors, including NSCLCs [38]. In our study, we also found the significant overexpression of PLAC1 (fold change > 90,  $P < 0.001$ ) in LSCC tissues. These findings above suggest that ENST00000441841 may



**Figure 5: ROC curve analysis.** ROC curves of the panel in training set (A) and validation set (B) for the discrimination of LSCCs from NT samples. AUC: area under the curve.

be involved in the activation of onco-placental gene PLAC1. Uc011cly.2 is transcribed from the sense strand of chromosome 5 and is overlapping with the transcript of gene SLC9A3 (solute carrier family 9 member a3). As a nature antisense lncRNA [25, 39], uc011cly.2 may be involved in the regulation of target gene SLC9A3 and its product. In summary, the associated genes mentioned above may offer clues for the function study of these five lncRNA biomarkers, however, further studies are still needed to illuminate their roles in the cancer processes. Collectively, our study is the first study that evaluate the potential of lncRNAs as biomarkers in early stage LSCC, and also provide important resource for further biological researches.

## MATERIALS AND METHODS

### Study samples

Paired LSCC and adjacent NT samples were collected from 63 patients (Table 1) who underwent primary surgical resection between August 2014 and September 2015 at Shanghai Zhongshan Hospital (Shanghai, China). The early stage LSCC in our study is defined as LSCC before stage III. No tumor metastasis was observed at the time of resection. All patients gave informed consent prior to sample collection. The pathological stage, grade, and nodal status was evaluated by an experienced pathologist and her colleagues. The patient's stage was determined according to the 7th edition of the TNM Classification of Malignant Tumors [40]. This study was approved by the Institutional Review Board of Shanghai Zhongshan Hospital.

### RNA extraction

Paired LSCC and adjacent NT samples from each subject were snap-frozen in liquid nitrogen immediately after resection and stored at  $-80^{\circ}\text{C}$  until later use. Total RNA was isolated from the frozen tissues sample using All repo DNA/RNA Mini Kit (QIAGEN, Valencia, CA, USA) according to the manufacturer's instructions. Total RNA from each sample was quantified using a NanoDrop ND-1000 (Nano-Drop Technologies, Wilmington, DE, USA), and RNA integrity was assessed using standard denatured agarose gel electrophoresis.

### lncRNA and mRNA microarray expression profiling

Three pairs of early stage LSCC samples from 63 total were selected for lncRNA microarray analysis using an Arraystar Human lncRNA Microarray V3.0 (Arraystar, Rockville, MD, USA). The GEO accession number is GSE88862. Sample preparation and microarray

hybridization was performed according to the manufacturer's standard protocols with minor modifications (Supplementary Method). After the hybridization and array image acquisition, quantile normalization and subsequent data processing were performed using the GeneSpring GX v11.5.1 software package (Agilent Technologies, Santa Clara, CA, USA). The  $p$ -values were adjusted through Benjamini-Hochberg method. The threshold used to determine up- or down-regulated lncRNAs and mRNAs was a fold-change  $> 2.0$  and FDR  $< 0.05$ . All microarray work was performed by Kangchen Bio-Tec (Shanghai, P.R. China).

### Co-expression and functional group analysis

Co-expression analysis was performed by associating the expression profiles of DE-lncRNAs with DE-mRNAs [41]. This association was based on the information of the lncRNAs' associated coding regions, which was supplied by microarray analysis. GO analysis was derived from Gene Ontology (<http://www.geneontology.org>). We used GO analysis to associate differentially expressed mRNA with GO functional categories. We also performed a pathway analysis for the differentially expressed mRNA based on the latest KEGG (Kyoto Encyclopedia of Genes and Genomes) database. The  $p$ -value (Hypergeometric- $P$  value) denoted the significance of the pathway correlated to the conditions. Recommend  $p$ -value cut-off (0.05) was used.

### Selecting candidate biomarkers

lncRNA was chosen for further selection if its raw expression intensity was between 100 and 20,000 in all three pairs of LSCC samples. The average raw expression intensity across all three tumor samples ( $\tilde{T}$ ) and nontumorous samples ( $\tilde{N}$ ) was then calculated. The difference between  $\tilde{T}$  and  $\tilde{N}$  was denoted as  $A=|\tilde{T}-\tilde{N}|$ .  $N_{\max}$  and  $N_{\min}$  represents the maximum and minimum raw intensity in NT samples respectively;  $T_{\max}$  and  $T_{\min}$  represents the maximum and the minimum raw intensity in LSCC samples respectively. A given lncRNA was selected as a candidate biomarker if either  $N_{\max} - N_{\min}$  or  $T_{\max} - T_{\min}$  was smaller than  $A/10$ .

### Quantitative PCR

Complementary DNA (cDNA) was reverse transcribed and synthesized using the SuperScript III First Strand Synthesis System (Invitrogen, Carlsbad, CA, USA) according to the manufacturer's instructions. We added approximately 1  $\mu\text{g}$  of total RNA to each reaction. The RT-qPCR was performed on a Light Cycler 480 (Roche Diagnostics, Penzberg, Germany) using LightCycler 480 SYBRGreen I Master (Roche Diagnostics, Mannheim, Germany). The thermocycling protocol included the following steps: denaturation step

of 10 min at 95°C; 40 cycles of 95°C for 15 s, and 60°C for 1 min. after the amplification, a melting curve analysis was performed at 95°C for 5 s and 65°C for 1 min to monitor primer dimers or nonspecific product formation. Each sample was measured in triplicate and a negative control was included with every RT-qPCR assay.  $\beta$ -actin was used as the reference gene instead of GAPDH since a recent study reported that the latter had differential expression between LSCC and NT samples [42]. Relative expression levels were quantified based on the crossing point values and normalized to the reference gene. The gene expression levels were calculated by the  $2^{-\Delta\Delta C_p}$  method. The primer sequences used in the RT-qPCR are listed in Supplementary Table 7.

### Statistical analysis

SPSS version 22.0 software (SPSS, Inc., Chicago, IL) was used for all statistical analyses. Student's *t*-test was used to analyze the lncRNA expression differences between samples of LSCC and adjacent NT samples. ROC curves were constructed to discriminate between the two different kinds of samples. The AUC value was used to assess for predictive power. The sensitivity and specificity were calculated according to the standard formulas.

### ACKNOWLEDGMENTS

We thank the Shanghai Zhongshan hospital (Shanghai, China) for providing samples and clinical information of early stage LSCC patients.

### CONFLICTS OF INTEREST

The authors disclosed no potential conflicts of interest.

### GRANT SUPPORT

This work was supported by Grants from the National Science Foundation of China (No.61571429, No.61271162 and No.61302052), the Science and Technology Commission of Shanghai Municipality (16410711800 and 15DZ1940105) and the Council of Shanghai Municipal Government for Science and Technology, the Ministry of Health of China for Clinical Key Discipline in 2012.

### REFERENCES

1. Siegel R, Ma J, Zou Z, Jemal A. Cancer statistics, 2014. *CA Cancer J Clin.* 2014; 64:9–29.
2. Davidson MR, Gazdar AF, Clarke BE. The pivotal role of pathology in the management of lung cancer. *J Thorac Dis.* 2013; 5:S463–78.

3. Kulpa J, Wójcik E, Reinfuss M, Kołodziejcki L. Carcinoembryonic antigen, squamous cell carcinoma antigen, CYFRA 21-1, and neuron-specific enolase in squamous cell lung cancer patients. *Clin Chem.* 2002; 48:1931–7.
4. Wilusz JE, Sunwoo H, Spector DL. Long noncoding RNAs: functional surprises from the RNA world. *Genes Dev.* 2009; 23:1494–504.
5. Sahu A, Singhal U, Chinnaiyan AM. Long noncoding RNAs in cancer: from function to translation. *Trends in Cancer.* 2015.
6. Lee JT. Epigenetic regulation by long noncoding RNAs. *Science.* 2012; 338:1435–9.
7. Yan S, Yingyi W, Wenkang L, Ping W, Tao T, Junxia Z, Jin Q, Ning L, Yongping Y. Long Non-Coding RNA H19 Promotes Glioma Cell Invasion by Deriving miR-675. *PLoS One.* 2014; 9:e86295.
8. Flynn RA, Chang HY. Long noncoding RNAs in cell-fate programming and reprogramming. *Cell Stem Cell.* 2014; 14:752–61.
9. Huarte M, Guttman M, Feldser D, Garber M, Koziol MJ, Kenzelmann-Broz D, Khalil AM, Zuk O, Amit I, Rabani M. A large intergenic noncoding RNA induced by p53 mediates global gene repression in the p53 response. *Cell.* 2010; 142:409–419.
10. Jie L, Liang T, Jie Z, Jian N, Hai-Ping Z, Ling Z, Jian-Fang X, Di Z. Long non-coding RNA CARLo-5 is a negative prognostic factor and exhibits tumor pro-oncogenic activity in non-small cell lung cancer. *Tumor Biol.* 2014; 35:11541–9.
11. Qiu M, Xu Y, Yang X, Wang J, Hu J, Xu L, Yin R. CCAT2 is a lung adenocarcinoma-specific long non-coding RNA and promotes invasion of non-small cell lung cancer. *Tumor Biol.* 2014; 35:5375–80.
12. Hou Z, Zhao W, Zhou J, Shen L, Zhan P, Xu C, Chang C, Bi H, Zou J, Yao X, Huang R, Yu L, Yan J. A long noncoding RNA Sox2ot regulates lung cancer cell proliferation and is a prognostic indicator of poor survival. *Int J Biochem Cell Biol.* 2014; 53:380–8. doi: 10.1016/j.biocel.2014.06.004.
13. Kai-Hua L, Wei L, Xiang-Hua L, Ming S, Mei-Ling Z, Wei-Qin W, Wei-Ping X, Ya-Yi H. Long non-coding RNA MEG3 inhibits NSCLC cells proliferation and induces apoptosis by affecting p53 expression. *BMC Cancer.* 2013; 13:4989–94.
14. Zhang EB, Yin DD, Sun M, Kong R, Liu XH, You LH, Han L, Xia R, Wang KM, Yang JS. P53-regulated long non-coding RNA TUG1 affects cell proliferation in human non-small cell lung cancer, partly through epigenetically regulating HOXB7 expression. *Cell Death Dis.* 2014; 5:e1243–e1243.
15. Qiu M, Xu Y, Wang J, Zhang E, Sun M, Zheng Y, Li M, Xia W, Feng D, Yin R. A novel lncRNA, LUADT1, promotes lung adenocarcinoma proliferation via the epigenetic suppression of p27. *Cell Death Dis.* 2015;6.
16. Wang P, Lu S, Mao H, Bai Y, Ma T, Cheng Z, Zhang H, Jin Q, Zhao J, Mao H. Identification of biomarkers for the detection of early stage lung adenocarcinoma by microarray profiling of long noncoding RNAs. *Lung Cancer.* 2015; 88:147–53.



17. Guenzl PM, Barlow DP. Macro lncRNAs A new layer of cis-regulatory information in the mammalian genome. *RNA Biol.* 2012; 9:731–41.
18. Lina M, Vladimir B B, Zhang Z. On the classification of long non-coding RNAs. *RNA Biol.* 2013; 10:924–33.
19. Dj. S. Wnt Signaling Pathway in Non-Small Cell Lung Cancer. *Jnci J Natl Cancer Inst.* 2014;106: djt356-djt356.
20. Jian L, Cho SN, Akkanti B, Jin N, Mao J, Long W, Chen T, Zhang Y, Tang X, Wistub I. ErbB2 Pathway Activation upon Smad4 Loss Promotes Lung Tumor Growth and Metastasis. *Cell Rep.* 2015; 10:1599–1613.
21. Xiaogang T, Wenyan Q, Liang Z, Jie H, Baozhong L, Cuiyan Z, Junting W, Fang Z, Kang S, Yimin S. A 5-microRNA signature for lung squamous cell carcinoma diagnosis and hsa-miR-31 for prognosis. *Clin Cancer Res An Off J Am Assoc Cancer Res.* 2011; 17:6802–11.
22. Mercer TR, Dinger ME, Mattick JS. Long non-coding RNAs: insights into functions. *Nat Rev Genet.* 2009; 10:155–9. doi: 10.1038/nrg2521.
23. Ponting CP, Oliver PL, Reik W. Evolution and Functions of Long Noncoding RNAs. *Cell.* 2009; 629–41. doi: 10.1016/j.cell.2009.02.006.
24. Huang B, Song JH, Cheng Y, Abraham JM, Ibrahim S, Sun Z, Ke X. Long non-coding antisense RNA KRT7-AS is activated in gastric cancers and supports cancer cell progression by increasing KRT7 expression. *Oncogene.* 2016; 35:4927–36.
25. Schein A, Zucchelli S, Kauppinen S, Gustincich S, Carninci P. Identification of antisense long noncoding RNAs that function as SINEUPs in human cells. *Sci Rep.* 2016; 6.
26. Su X, Malouf GG, Chen Y, Zhang J, Yao H, Valero V, Weinstein JN, Spano JP, Meric-Bernstam F, Khayat D. Comprehensive analysis of long non-coding RNAs in human breast cancer clinical subtypes. *Oncotarget.* 2014; 5:9864. doi: 10.18632/oncotarget.2454.
27. Tang J, Jiang R, Deng L, Zhang X, Wang K, Sun B. Circulation long non-coding RNAs act as biomarkers for predicting tumorigenesis and metastasis in hepatocellular carcinoma. *Oncotarget.* 2015; 6:4505. doi:10.18632/oncotarget.2934.
28. Tinzl M, Marberger M, Horvath S, Chypre C. DD3 PCA3 RNA analysis in urine—a new perspective for detecting prostate cancer. *Eur Urol.* 2004; 46:182–7.
29. Chen CH, Wang SW, Chen CW, Huang MR, Hung JS, Huang HC, Lin HH, Chen RJ, Shyu MK, Huang MC. MUC20 overexpression predicts poor prognosis and enhances EGF-induced malignant phenotypes via activation of the EGFR–STAT3 pathway in endometrial cancer. *Gynecol Oncol.* 2013; 128:560–7.
30. Arseneault R, Chien A, Newington JT, Rappon T, Harris R, Cumming RC. Attenuation of LDHA expression in cancer cells leads to redox-dependent alterations in cytoskeletal structure and cell migration. *Cancer Lett.* 2013; 338:255–66.
31. Sheng SL, Liu JJ, Dai YH, Sun XG, Xiong XP, Huang G. Knockdown of lactate dehydrogenase A suppresses tumor growth and metastasis of human hepatocellular carcinoma. *FEBS J.* 2012; 279:3898–910.
32. Yao F, Zhao T, Zhong C, Zhu J, Zhao H. LDHA is necessary for the tumorigenicity of esophageal squamous cell carcinoma. *Tumor Biol.* 2013; 34:25–31.
33. Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Fan L, Zytnicki M, Notredame C, Huang Q. Long Noncoding RNAs with Enhancer-like Function in Human Cells. *Cell.* 2010; 143:46–58.
34. Du Z, Fei T, Verhaak RGW, Su Z, Zhang Y, Brown M, Chen Y, Liu XS. Integrative genomic analyses reveal clinically relevant long noncoding RNAs in human cancer. *Nat Struct Mol Biol.* 2013; 20:908–13.
35. Yang J, Lin J, Liu T, Chen T, Pan S, Huang W, Li S. Analysis of lncRNA expression profiles in non-small cell lung cancers (NSCLC) and their clinical subtypes. *Lung Cancer.* 2014;110–5. doi: 10.1016/j.lungcan.2014.05.011.
36. Wan L, Zhang L, Fan K, Cheng ZX, Sun QC, Wang JJ. Knockdown of Long Noncoding RNA PCAT6 Inhibits Proliferation and Invasion in Lung Cancer Cells. *Oncol Res.* 2016; 24:161–70.
37. Chang W, Yang Q, Zhang H, Lin H, Zhou Z, Lu X, Zhu C, Xue LQ, Wang H. Placenta-specific protein 1 (PLAC1) is involved in the trophoblasts invasion and migration. *Reproduction.* 2014;148.
38. Jr SW, Gnjatic S, Ritter E, Chua R, Cohen T, Hsu M, Jungbluth AA, Altorki NK, Chen YT, Old LJ. PLAC1, a trophoblast-specific cell surface protein, is expressed in a range of human tumors and elicits spontaneous antibody responses. *Cancer Immun.* 2007; 7:18.
39. Zucchelli S, Fasolo F, Russo R, Cimatti L, Patrucco L, Takahashi H, Jones MH, Santoro C, Sblattero D, Cotella D, Persichetti F, Carninci P, Gustincich S. SINEUPs are modular antisense long non-coding RNAs that increase synthesis of target proteins in cells. *Front Cell Neurosci.* 2015; 9:174. doi: 10.3389/fncel.2015.00174.
40. Goldstraw P, Crowley J, Chansky K, Giroux DJ, Groome PA, Rami-Porta R, Postmus PE, Rusch V, Sobin L, Committee IA for the S of LCIS. The IASLC Lung Cancer Staging Project: proposals for the revision of the TNM stage groupings in the forthcoming (seventh) edition of the TNM Classification of malignant tumours. *J Thorac Oncol.* 2007; 2:706–14.
41. Longxin W, Dian F, Yongbin Q, Xiaoxiao X, Feng X, Conghui H, Xiaofeng X, Zhifeng W, Zhengyu Z, Jingping G. Genome-wide screening and identification of long noncoding RNAs and their interaction with protein coding RNAs in bladder urothelial cell carcinoma. *Cancer Lett.* 2014; 349:77–86.
42. Zhan C, Zhang Y, Ma J, Wang L, Jiang W, Shi Y, Wang Q. Identification of reference genes for qRT-PCR in human lung squamous-cell carcinoma by RNA-Seq. *Acta Biochim Biophys Sin (Shanghai).* 2014; 46:330–7.