

# Differential expression profiles of microRNAs as potential biomarkers for the early diagnosis of lung cancer

YANQIU ZHANG, JING SUI, XIAN SHEN, CHENGYUN LI, WENZHUO YAO, WEIWEI HONG,  
HUI PENG, YUEPU PU, LIHONG YIN and GEYU LIANG

Key Laboratory of Environmental Medicine Engineering, Ministry of Education,  
School of Public Health, Southeast University, Nanjing, Jiangsu 210009, P.R. China

Received October 14, 2016; Accepted April 10, 2017

DOI: 10.3892/or.2017.5612

**Abstract.** Lung cancer is one of the most lethal malignancies worldwide. To reduce the high morbidity and mortality of the disease, sensitive and specific biomarkers for early detection are urgently needed. Tumor-specific microRNAs (miRNAs) seem to be potential biomarkers for the early diagnosis and treatment of cancer. In this study, the microarray of miRNAs and mRNAs on the same samples was performed and the intersection taken with The Cancer Genome Atlas (TCGA) lung cancer miRNA/RNAseq dataset. Then, miRNA-mRNA regulatory network was constructed to identify miRNA candidates associated with lung cancer through integrating gene expression and miRNA-target prediction. Furthermore, the expression levels of miRNA candidates were validated by stem-loop real-time reverse transcription PCR (qRT-PCR) in larger lung cancer population. The relationship between signature miRNAs and the risk of lung cancer were assessed by conditional logistic regression analysis. Diagnostic value of these miRNAs was determined by areas under receiver operating characteristic curves (ROC). The Affymetrix microarray analysis identified a total of 116 miRNAs and 502 mRNAs that could distinguish lung tumor tissues from adjacent non-tumor tissues, of which 70 miRNAs and 136 mRNAs were upregulated, while 46 miRNAs and 366 mRNAs were downregulated, respectively. In combination with TCGA analysis, we identified 32 miRNAs and 377 mRNAs related to lung cancer. Then, 28 key miRNAs related to 61 intersection mRNAs were identified by miRNA-mRNA network analysis. The miRNA function analysis was indicative of that 18 upregulated and 10 downregulated miRNAs involved in signaling pathways related to Environmental Information

Processing and Human Diseases. Population result showed that the expression of 7 miRNAs (miR-205-5p, miR-3917, miR-30a-3p, miR-30a-5p, miR-30c-2-3p, miR-30d-5p and miR-27a-5p) was consistent with the analysis result of microarray and TCGA. In addition, upregulation of miR-205-5p, miR-3917 and downregulation of miR-30a-3p, miR-30a-5p, miR-30c-2-3p, miR-30d-5p, miR-27a-5p increased the risk of lung cancer by conditional logistic regression analysis. The diagnostic accuracy of miR-205-5p, miR-3917, miR-27a-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p, miR-30d-5p showed that their corresponding AUCs were 0.728, 0.661, 0.637, 0.758, 0.772, 0.734, 0.776, respectively. Therefore, there are a set of signature miRNAs which may be promising biomarkers for the early screening of high-risk populations and early diagnosis of lung cancer.

## Introduction

With increasing incidence and mortality, cancer is the leading cause of death in China and is a major public health problem. Based on GLOBOCAN estimates, approximately 14.1 million new cancer cases and 8.2 million deaths occurred in 2012 worldwide (1). Moreover, based on the estimates of National Central Cancer Registry (NCCR) of China, 4.292 million new cancer cases and 2.814 million cancer deaths would occur in China in 2015, lung cancer is the leading cause of cancer-related death worldwide (2), and the 5-year survival rate of patients with advanced lung cancer remains low.

miRNAs are a class of endogenous, single-stranded, small noncoding RNA molecules (~22 nucleotides), which are crucial for post-transcriptional regulators of gene expression (3,4). In recent years, many studies have reported promising biomarkers for differential diagnosis of lung cancer (5-8). However, accurate biomarkers of lung cancer early diagnosis still remain largely unexplored. The discovery of miRNAs, are frequently dysregulated in lung cancer and are implicated in lung cancer growth, recurrence, and metastasis (9,10). Consequently, miRNAs are considered as valuable biomarkers for diagnosis of lung cancer (11,12). Evidence from recent studies indicates promise of miRNAs in risk assessment, prevention, early diagnosis, and prognosis of lung cancer (9,13-16). miRNA expression profiling is becoming an increasingly important tool to reveal gene activity in cancer.

---

*Correspondence to:* Professor Geyu Liang, Key Laboratory of Environmental Medicine Engineering, Ministry of Education, School of Public Health, Southeast University, 87 Ding Jia Qiao, Nanjing, Jiangsu 210009, P.R. China  
E-mail: lianggeyu@163.com

**Key words:** lung cancer, microRNA, The Cancer Genome Atlas, expression profile

In this study, microarray of the miRNAs and mRNAs on the same samples was performed and the intersection taken with The Cancer Genome Atlas (TCGA) lung cancer miRNA/RNAseq dataset. Then, miRNA-mRNA regulatory network was constructed to identify miRNA candidates associated with lung cancer through integrating gene expression and miRNA-target prediction. Furthermore, the expression levels of miRNA candidates were validated by stem-loop real-time reverse transcription PCR (qRT-PCR) in larger lung cancer population. The relationship between signature miRNAs and the risk of lung cancer were assessed by conditional logistic regression analysis. Diagnostic value of these miRNAs was determined by areas under receiver operating characteristic curves (ROC). The aim of our study was to identify potential miRNA biomarkers of lung cancer for early screening of high-risk populations and early diagnosis.

## Materials and methods

**Specimen collection.** Three lung tumor tissues and their paired adjacent non-tumor tissues samples were collected for miRNA and mRNA microarray analysis. In addition, 81 patients, including 64 males and 17 females aged between 36-76 years, were recruited for qRT-PCR analysis from the Nanjing Chest Hospital, with their consent and agreement. All patients were confirmed by pathology as lung tumor without preoperative radiotherapy or chemotherapy. Lung tumor tissues and adjacent non-tumor tissues were obtained from surgical specimens immediately after resection from patients. The adjacent non-tumor tissues was located at least 5 cm away from tumor edge. Tissue samples (0.5 cm<sup>3</sup>) were immersed in RNA locker (Tianzhang, China) at 4°C overnight and stored at -20°C until use. This study was approved by the ethics committee of Zhongda Hospital Affiliated to Southeast University.

**Total RNA extraction.** Total RNA was extracted from both lung tumor tissues and adjacent non-tumor tissues using TRIzol reagent (Invitrogen, Carlsbad, CA, USA) according to the instructions. Concentration and integrity of extracted RNA was assessed using a NanoDrop 1000 spectrophotometer (NanoDrop Technologies, Houston, TX, USA). Finally, all RNA samples showing A260/280 ratios between 1.8 and 2.0 were further analyzed for RNA integrity using agarose gel electrophoresis. The quality and integrity of RNA were satisfactory for all samples.

**miRNA and mRNA microarray.** Affymetrix™ Genechip miRNA 4.0 Array (Affymetrix, Inc., Santa Clara, CA, USA) and Affymetrix Genechip HTA 2.0 Array (Affymetrix, Inc.) were used in this study. Samples of miRNAs and mRNAs from 6 specimens were labeled and hybridized according to the manufacturer's protocol. Signals were normalized using the median center tool for genes.

**TCGA analysis of differential expression of miRNAs and mRNAs.** To improve the reliability and accuracy of the studies, we analyzed miRNAs and mRNAs expression profile dataset of lung cancer from The Cancer Genome Atlas (TCGA) database. A total of 521 lung cancer patient were obtained from TCGA database. After criteria of exclusion, a total of

465 lung cancer patients were included in our study. In addition, the lung cancer RNA expression data (level 3) of the corresponding patients were downloaded from the TCGA Data Portal (March 2016). The RNA sequencing raw reads (mRNAs) were post-processed and normalized by TCGA RNASeqV2 system. Level 3, normalized miRNA expression data were downloaded from the TCGA Data Portal performed using Illumina HiSeq and Illumina GA miRNAs sequencing platforms (Illumina Inc., Hayward, CA, USA) and quantile normalized before performing analysis.

**Construction of miRNA-mRNAs network and bioinformatics analysis.** To construct miRNAs-RNAs network, the miRNA-mRNA interactions were download from the starBase v2.0 database (<http://starbase.sysu.edu.cn/>) (17), then aberrantly expressed intersection miRNAs and mRNAs with absolute value of fold change  $\geq 2$  and  $P < 0.05$  were retained. Targetscan (<http://www.targetscan.org/>) and miRTarBase (<http://mirtarbase.mbc.nctu.edu.tw/>) (18) were performed to predict miRNA target genes. The results could be verified in starBase database. Furthermore, in combination with microarray data and TCGA result, miRNA-mRNA co-expression network was constructed by Cytoscape v3.0 (19). The potential function of these miRNAs was analyzed by DIANA-mirPath v.3.

**miRNA expression analysis in lung cancer population by qRT-PCR.** To confirm the reliability and validity of the microarray data, we selected 7 (miR-205-5p, miR-3917, miR-27a-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p and miR-30d-5p) differentially expressed key miRNAs, which were found in non-small cell lung cancer pathway, and analyzed their actual expression levels in 81 pairs of diagnosed lung tumor and adjacent non-tumor tissues by qRT-PCR on StepOnePlus™ PCR System (Applied Biosystems, Waltham, MA, USA). Reverse transcription reactions were conducted in two steps. First, 0.5 µg of RNA samples was incubated in a 0.2 ml Ep tube for 5 min at 65°C and held at 4°C. Then, the 10 µl mixture which comprised 2 µl RT product which obtained from the front, 2 µl 5X RT buffer (Promega, Madison, WI, USA), 0.5 µl RT Enzyme Mix, 0.5 µl miRNA-specific stem-loop RT primers (RiboBio, Guangzhou, China), and 5 µl RNase-free water (Tiangen, China) was incubated in a 0.2 ml Ep tube for 15 min at 37°C, 5 min at 98°C and subsequently held at 4°C. PCR was then performed using Thunderbird SYBR qPCR Mix (Toyobo Corp., Osaka, Japan) according to the manufacturer's protocol. The PCR reaction components were 1 µl cDNA, 5 µl Thunderbird SYBR qPCR Mix, 0.3 µl PCR primers (RiboBio) and 3.4 µl RNase-free water. The reaction was performed at 95°C for 1 min, followed by 40 cycles of 95°C for 15 sec, 60°C for 30 sec and 72°C for 30 sec. Dissociation curve was analyzed from 60 to 90°C. U6 was chosen as the endogenous standard. All assays were performed in triplicates.

**Statistical analysis.** The microarray data and TCGA data analyses were performed using one-way ANOVA, and absolute value of fold change  $\geq 2$  and  $P < 0.05$  was considered to indicate statistically significant differences. qRT-PCR results were expressed as mean  $\pm$  SD. miRNA expression was calculated using the  $2^{-\Delta\Delta Ct}$  method, which  $\Delta Ct = Ct_{miRNAs} - Ct_{U6}$  and  $\Delta\Delta Ct = \Delta Ct_{tumor tissues} - \Delta Ct_{adjacent non-tumor tissues}$ . The threshold cycle

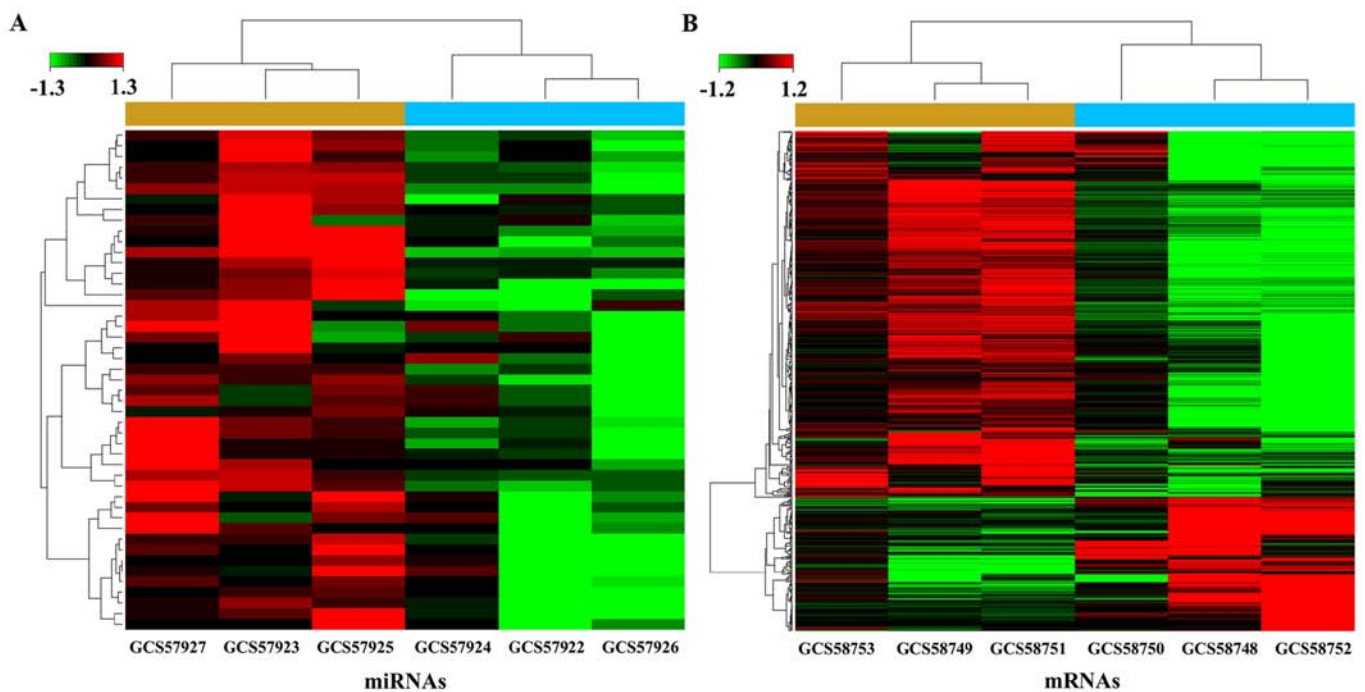


Figure 1. Cluster analysis of differentially expressed miRNAs (A) and mRNAs (B) in lung tumor tissues vs. adjacent non-tumor tissues.

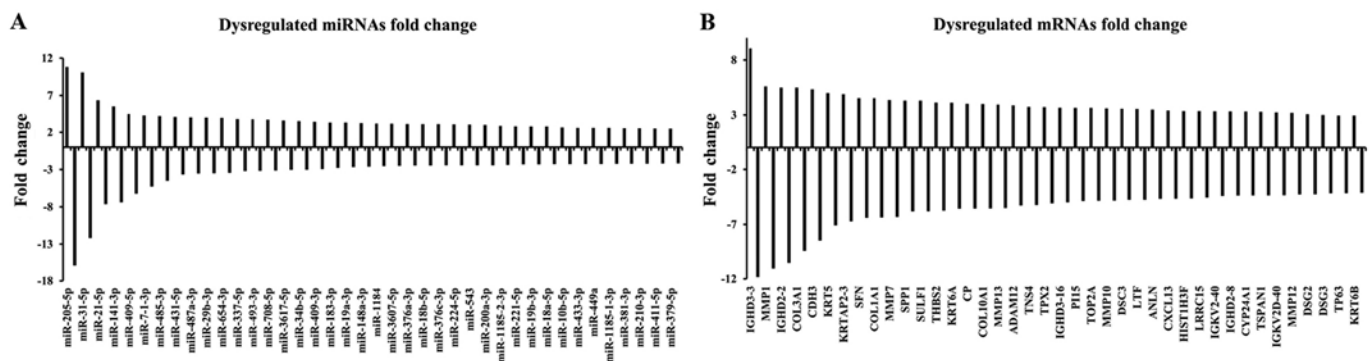


Figure 2. Top 40 differential expression levels of miRNAs (A) and mRNAs (B) in lung tumor tissues vs. adjacent non-tumor tissues by microarray analysis.

indicates the fractional cycle number at which the amount of amplified target reaches a fixed threshold and the Ct value is negatively correlated with copy numbers. Thus, in the final calculation, the Ct values are multiplied with -1 for logistic regression analysis. Statistical analysis was carried out with the Student's t-test for comparison of two groups in qRT-PCR analysis, and conditional logistic regression for evaluate the association between differentially expressed miRNAs and the risk of lung cancer. In both cases, differences with  $P < 0.05$  were considered to indicate a statistically significant result. Operating characteristic curve (ROC) was used to determine the signature miRNAs as the sensitivity and specificity of detection of lung cancer.

## Results

**miRNA and mRNA microarray analysis.** Microarray analysis identified 116 miRNAs and 502 mRNAs that could distinguish lung tumor tissues from adjacent non-tumor tissues

(fold change  $> 2$ ). A total of 70 miRNAs and 136 mRNAs were upregulated, while 46 miRNAs and 366 mRNAs were downregulated. Cluster analysis, based on the differentially expressed miRNAs and mRNAs, successfully separated the tumor tissues from the adjacent non-tumor tissues (Fig. 1). The analysis manifested the chips in each case consistently. Top 40 differentially expressed miRNAs and mRNAs are listed (Fig. 2). miR-205-5p and miR-31-5p (fold change 10.825 and 10.083, respectively) were the most upregulated miRNAs, and miR-4521 and miR-486-5p (fold change -15.948 and -12.172) were the most downregulated. IGH23-3 and MMP1 (fold change 9.050 and 5.598, respectively) were the most upregulated mRNAs, SLC6A4 and RTKN2 (fold change -11.842 and -11.063, respectively) were the most downregulated.

**TCGA analysis of differential expression of miRNAs and mRNAs.** We identified a total of 1030 miRNAs and 18633 mRNAs from TCGA database, of these 118 and 3853

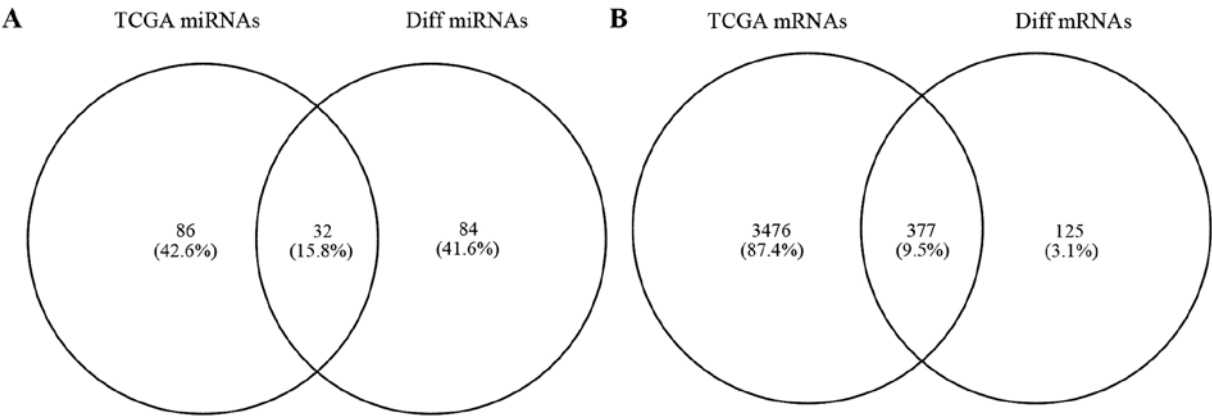


Figure 3. Venn diagram analysis of differentially expressed miRNAs (A) and mRNAs (B) between TCGA data and microarray data.

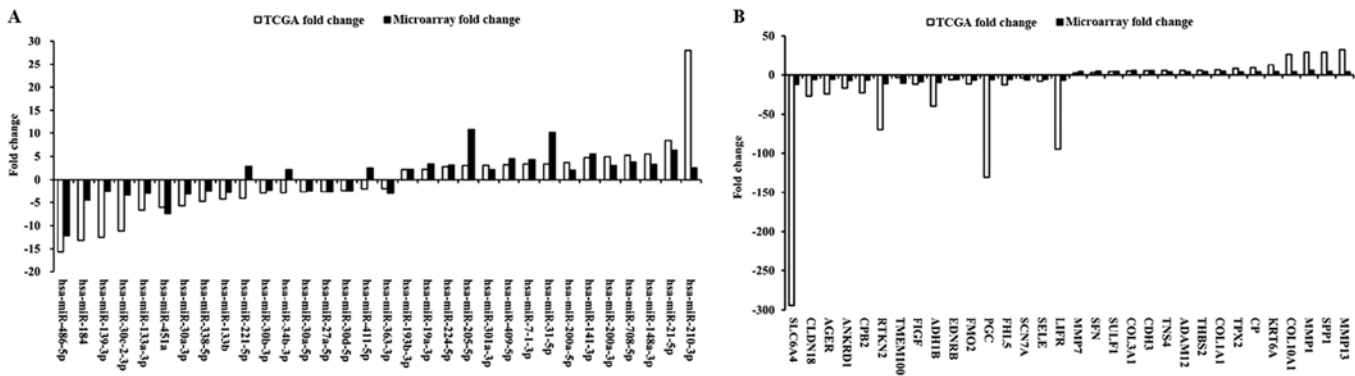


Figure 4. Top 32 differential expression miRNAs (A) and mRNAs (B) after take intersection analysis between TCGA and microarray data.

differentially expressed miRNAs (absolute fold change  $>2$ ,  $P<0.05$ ) and mRNAs (absolute fold change  $>2$ ,  $P<0.05$ ) were found between lung tumor tissues and adjacent non-tumor lung tissues. Based on these data, differentially expressed miRNAs and mRNAs were selected for further analysis. Then 32 and 377 differentially expressed miRNAs and mRNAs were obtained for further analysis after intersection between TCGA and microarray data (Figs. 3 and 4).

**miRNA target prediction and miRNA-mRNA network construction.** In order to establish miRNA-mRNA network, mRNAs targeted by miRNAs was performed. Based on the miRNAs and mRNAs described in Figs. 3 and 4, starBase v2.0 database, TargetScan and miRTarbase were performed to predict miRNAs-targeted mRNAs. Then, we identified 28 specific miRNAs related to 61 intersection mRNAs (Table I). As described in Table I, the miRNA-mRNA network was constructed via Cytoscape 3.0. Fig. 5, shows the 28 miRNAs and 61 mRNAs involved in the proposed network.

**Function analysis of differentially expressed miRNAs.** We next used the DIANA-mirPath V.3, which the Kyoto Encyclopedia of Genes and Genomes (KEGG) to classify and analyze the potential functions of these 28 miRNAs in the pathways. Our analysis indicated that 18 upregulated and 10 downregulated miRNAs were involved in signaling pathways related to

Environmental Information Processing and Human Diseases, with  $-\ln(P\text{-value}) >3$  (Fig. 6). The former group included the Hippo, TGF- $\beta$ , Wnt, ErbB and MAPK signaling pathway corresponded to up-/downregulated miRNAs, respectively. While the latter group included proteoglycans in cancer, renal cell carcinoma, glioma, pathways in cancer and non-small cell lung cancer corresponded to up-/downregulated miRNAs, respectively. Moreover, 23 miRNAs and 29 gene involved in signaling pathways related to non-small cell lung cancer pathway, respectively.

**miRNA expression analysis in lung cancer population by qRT-PCR.** qRT-PCR results demonstrated that expression of miR-205-5p and miR-3917 were significantly higher in tumor tissues than adjacent non-tumor tissues ( $P<0.05$ ), while expression of miR-27a-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p, and miR-30d-5p were significantly lower in tumor tissues than adjacent non-tumor tissues ( $P<0.05$ ). Expression levels of these 7 miRNAs were consistent with the microarray results and TCGA data (Table II; Figs. 7 and 8).

**Association between verified miRNAs and lung cancer.** Conditional logistic regression analysis was used to evaluate the association between differentially expressed miRNAs and the risk of lung cancer. As shown in Table III, significantly increased risk for lung cancer was associated with increased

Table I. Differential expression miRNAs and their targeted mRNA prediction.

miRNAs	miRNA feature	Gene symbol	Gene feature
miR-133a-3p	Down	COL1A1, THBS2	Up
miR-133b	Down	MMP9	Up
miR-184	Down	SLC7A5	Up
miR-30a-5p	Down	CDC20, ITGA2, KRT6B, SERPINB5, SLC7A5	Up
miR-30b-3p	Down	FAM83B, HIST1H2BD, SLC7A5, TNS4	Up
miR-30c-2-3p	Down	BUB1, FAM83B, TNS4	Up
miR-30d-5p	Down	SLC7A5	Up
miR-338-5p	Down	CLDN1	Up
miR-363-3p	Down	ARNTL2, HMGA2	Up
miR-451a	Down	MMP9	Up
miR-141-3p	Up	GATA6	Down
miR-148a-3p	Up	S1PR1	Down
miR-193b-3p	Up	TGFBR3, TMTC1, RTKN2, TACC1	Down
miR-19a-3p	Up	NR4A2, LDLR, ARHGEF26, SOCS3, PTPRB, SFTPA1, TMTC1, HECW2	Down
miR-200a-3p	Up	GATA6, DLC1	Down
miR-200a-5p	Up	HHIP	Down
miR-205-5p	Up	LRRK2	Down
miR-210-3p	Up	SERTM1	Down
miR-21-5p	Up	TIMP3, BTG2, IL1B, TCF21, EDIL3, CYBRD1, SASH1, FAXDC2, LIFR, SEMA5A, LIMCH1, OLR1, KLF9, WWC2, HPGD	Down
miR-221-5p	Up	LRRTM2, CLIC5	Down
miR-224-5p	Up	FOSB	Down
miR-301a-3p	Up	CYBRD1, LDLR, ARHGEF26, EDN1, SFTPA1, DLC1, MYH11, TMTC1	Down
miR-31-5p	Up	SELE	Down
miR-34b-3p	Up	ACADL, ADH1B, PAQR5, PLCE1	Down
miR-409-5p	Up	ABCA3	Down
miR-411-5p	Up	ART4, ITM2A	Down
miR-708-5p	Up	CD274, PAQR5	Down
miR-7-1-3p	Up	LDLR, LIFR	Down

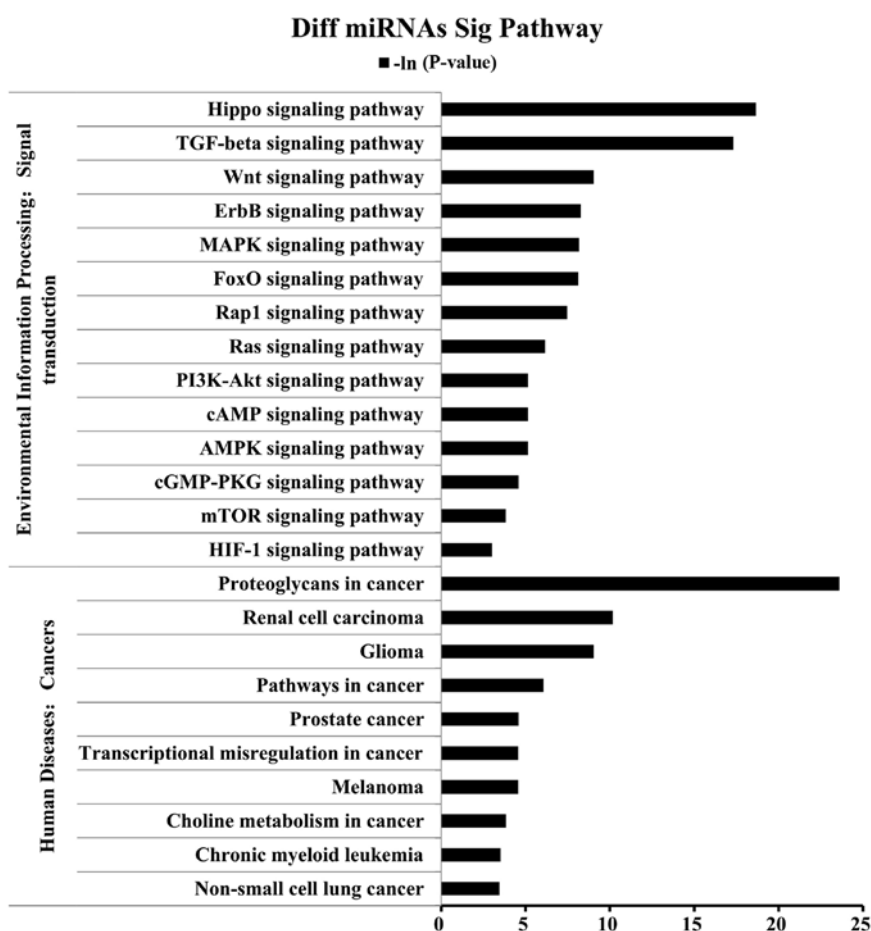
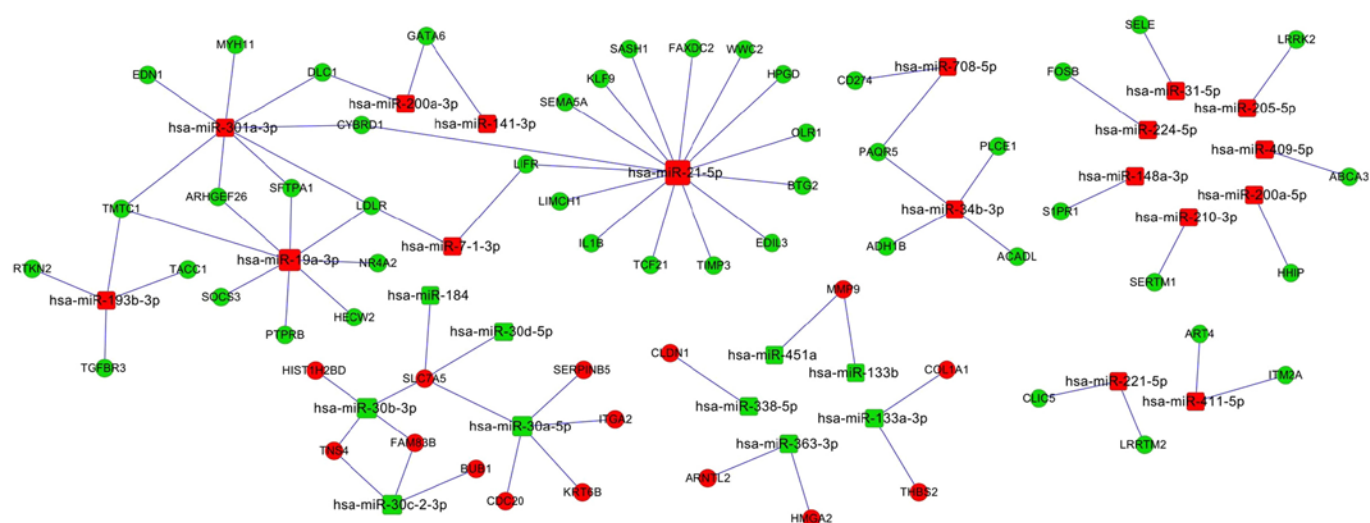
expression of miR-205-5p and miR-3917 (OR 1.287 and 1.360, respectively) and reduced expression of miR-27a-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p and miR-30d-5p (OR 0.846, 0.700, 0.585, 0.745 and 0.479, respectively). This suggested that these miRNAs could impact on the risk of lung cancer and play an important role during lung cancer development.

**ROC curve analysis of signature miRNAs.** ROC curve analysis demonstrated the area under curve (AUC) 0.728, 0.758, 0.772, 0.734, 0.776 for miR-205-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p and miR-30d-5p (P-value <0.01, 95% confidence interval: 0.650-0.806, 0.683-0.832, 0.699-0.845, 0.650-0.817 and 0.704-0.848, respectively, Table IV and Fig. 9A, D-G), show the higher cutoff (0.7) and could be considerable biomarker for early diagnosis of lung cancer. ROC analysis measured an AUC 0.661, 0.637 for miR-3917 and miR-27a-5p (P-value <0.01, 95% confidence interval: 0.578-0.744 and 0.549-0.725, respectively, Table IV and Fig. 9B and C),

show the score very close to the cutoff (0.7) needed for a considerable biomarker for early diagnosis of lung cancer. Furthermore, as total 7 miRNAs ROC curve conjoint analysis result, AUC=0.919 (P-value <0.01, 95% confidence interval: 0.874-0.963, Table IV and Fig. 9H), which is the score higher the cutoff (0.7) and could be considerable biomarkers for early diagnosis of lung cancer.

## Discussion

A growing body of evidence indicates that miRNAs are valuable source of markers for diagnosis in cancer (7,12,20-23). Chen *et al* (24) identified 10-serum miRNAs as potential markers for early diagnosis of lung cancer in 400 lung cancer cases and 220 controls. Yin *et al* found that increased expression of miR-150 is associated with poor prognosis in lung cancer by qRT-PCR (25). The above studies provide sound evidence that the alteration in the miRNA profiles could be

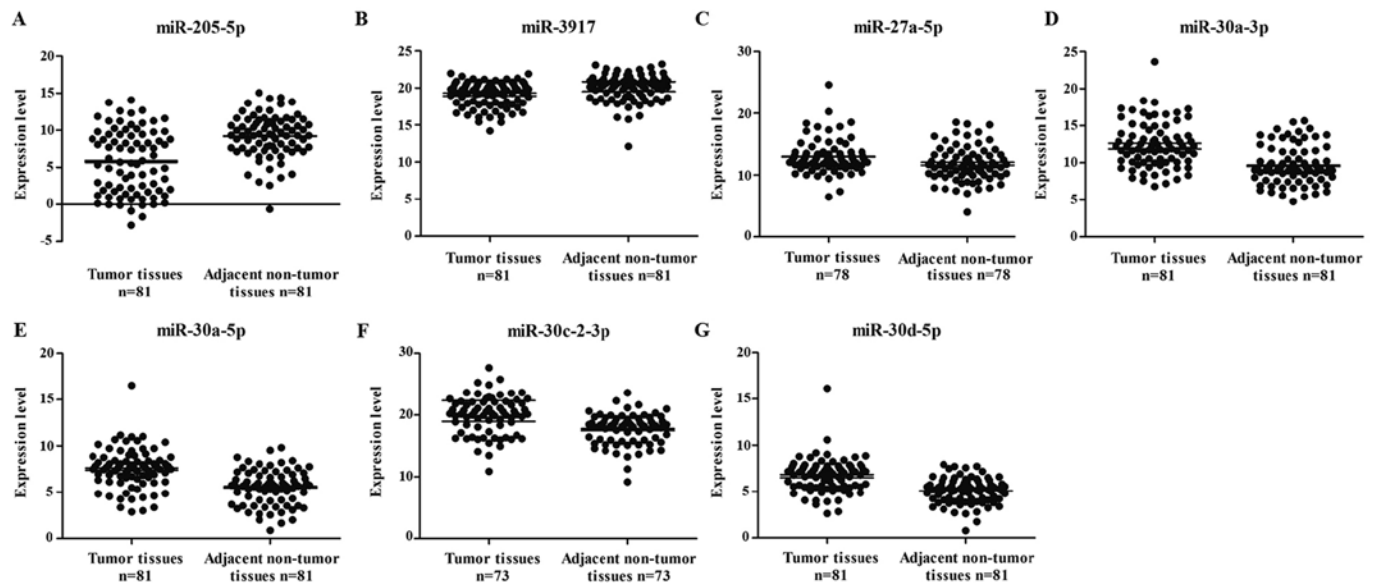


microarray techniques and miR-19b and miR-146a are potential biomarkers for the prediction of survival and response to chemotherapy in lung cancer using qRT-PCR in 94 healthy subjects and 94 advanced lung cancer patients. Molina-Pinelo *et al* (27) analyzed the expression of 381 miRNAs in bronchoalveolar lavage fluid samples from 87 patients classi-

Table II. Expression level of miRNAs analyzed by qRT-PCR in lung tumor tissues and non-tumor tissues.

miRNAs	N	Group <sup>a</sup>	$\Delta Ct^b$ (mean $\pm$ SD)	$\Delta\Delta Ct^b$ (mean $\pm$ SD)	$2^{-\Delta\Delta Ct}$	Fold change	P-value	t-value
miR-205-5p	81	Ca	5.757 $\pm$ 4.287	-3.423 $\pm$ 4.962	10.729	10.729	<0.001	-6.210
		N	9.181 $\pm$ 2.873					
miR-3917	81	Ca	19.047 $\pm$ 1.718	-0.937 $\pm$ 2.021	1.914	1.914	<0.001	-4.173
		N	19.984 $\pm$ 1.828					
miR-27a-5p	78	Ca	12.954 $\pm$ 2.724	1.206 $\pm$ 1.060	0.433	-2.309	<0.001	5.172
		N	11.748 $\pm$ 2.778					
miR-30a-3p	81	Ca	12.348 $\pm$ 3.059	2.718 $\pm$ 1.939	0.152	-6.579	<0.001	12.615
		N	9.630 $\pm$ 2.550					
miR-30a-5p	81	Ca	7.535 $\pm$ 2.168	2.019 $\pm$ 1.756	0.247	-4.049	<0.001	10.348
		N	5.516 $\pm$ 1.849					
miR-30c-2-3p	73	Ca	19.931 $\pm$ 3.084	2.238 $\pm$ 2.245	0.212	-4.717	<0.001	8.517
		N	17.692 $\pm$ 2.506					
miR-30d-5p	81	Ca	6.712 $\pm$ 1.846	1.640 $\pm$ 1.936	0.321	-3.115	<0.001	7.625
		N	5.072 $\pm$ 1.369					

<sup>a</sup>Ca: tumor tissues; N: adjacent non-tumor tissues. <sup>b</sup> $\Delta Ct = Ct_{miRNAs} - Ct_{u6}$ ;  $\Delta\Delta Ct = \Delta Ct_{tumor\ tissues} - \Delta Ct_{adjacent\ non-tumor\ tissues}$ .

Figure 7. Expression level of miRNAs analyzed ( $\Delta Ct$ ) by qRT-PCR in lung tumor tissues and adjacent non-tumor tissues.

fied into four groups: COPD, adenocarcinoma, adenocarcinoma with COPD, and control (neither COPD nor adenocarcinoma), and found that four miRNA clusters (the miR-17-92 cluster and its paralogues, miR-106a-363 and miR-106b-25, miR-192-194 and miR-132-212 cluster) were involved in the pathogenesis of adenocarcinoma and COPD, as biomarkers for diagnosis and/or therapeutic purposes. These reports indicate that specific miRNA profiles of lung cancer may bring breakthroughs to the diagnosis and treatment of lung cancer. However, the problem is that the results of these reports are not consistent

because of the small number of samples. Studies with large scale samples and microarray detection are rare, and the relationship between specific miRNAs and lung cancer still need further investigation. TCGA database is a public platform, in which large scale sequencing data of miRNA and mRNA for lung cancer can be downloaded. To improve the reliability and accuracy of the studies, we integrated the data from our miRNA microarray and TCGA database.

In the present study, we focused primary on whether promising miRNAs could act as biomarkers to discriminate lung

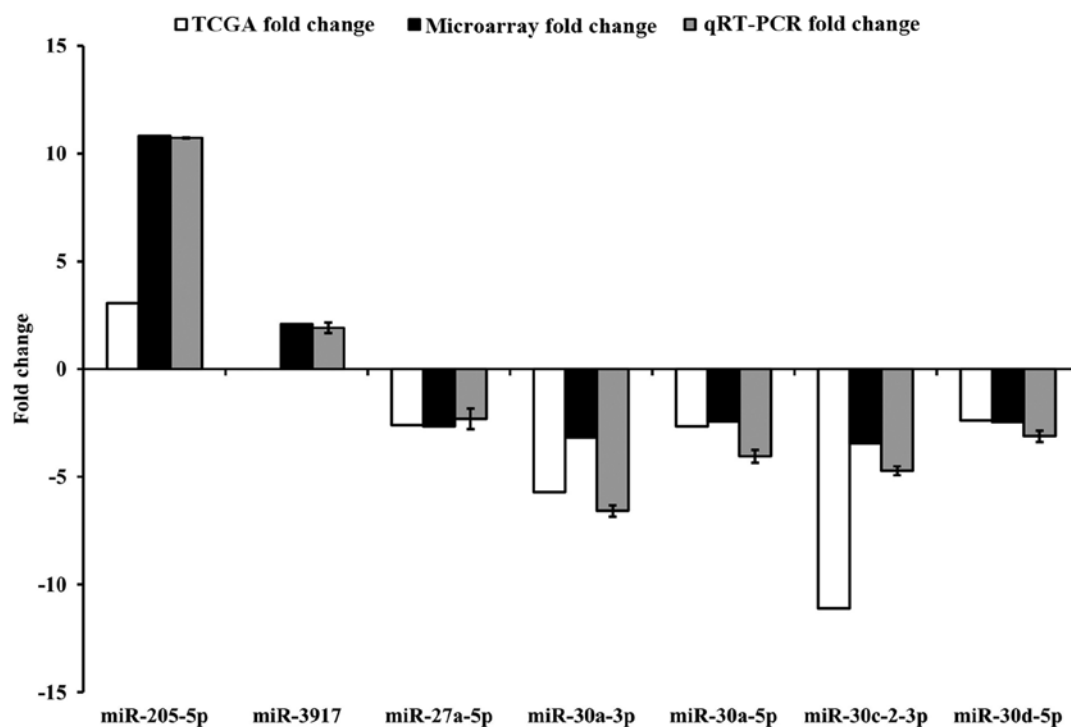


Figure 8. miRNAs expression level of miRNA microarray, TCGA and qRT-PCR.

Table III. Differentially expression of miRNAs associated with a high risk of lung cancer.

miRNAs	N	Group <sup>a</sup>	$\beta$	SE	Wald	P-value	OR	95% CI
miR-205-5p	81	Ca	0.252	0.051	24.887	<0.001	1.287	1.166-1.421
		N					1.000	
miR-3917	81	Ca	0.308	0.098	9.916	0.002	1.360	1.123-1.647
		N					1.000	
miR-27a-5p	78	Ca	0.167	0.064	6.756	0.009	0.846	0.746-0.960
		N					1.000	
miR-30a-3p	81	Ca	-0.357	0.070	25.815	<0.001	0.700	0.610-0.803
		N					1.000	
miR-30a-5p	81	Ca	-0.536	0.103	26.963	<0.001	0.585	0.478-0.716
		N					1.000	
miR-30c-2-3p	73	Ca	-0.294	0.070	11.475	<0.001	0.745	0.649-0.855
		N					1.000	
miR-30d-5p	81	Ca	-0.736	0.137	28.712	<0.001	0.479	0.366-0.627
		N					1.000	

<sup>a</sup>Ca: tumor tissues; N: adjacent non-tumor tissues.

cancer from normal cases by taking advantage of miRNA microarrays in lung tumor tissues and adjacent non-tumor tissues. A total of 116 miRNAs and 502 mRNAs were found dysregulated in tumor tissues. After comprehensive analysis between TCGA and microarray data, 32 and 377 differentially expressed miRNAs and mRNAs, which closely associated with lung cancer, were obtained. Furthermore, after miRNA-mRNA network construction, we identified 28

(miR-133a-3p, miR-133b, miR-184, miR-30a-5p, miR-30b-3p, miR-30c-2-3p, miR-30d-5p, miR-338-5p, miR-363-3p, miR-451a, miR-141-3p, miR-148a-3p, miR-193b-3p, miR-19a-3p, miR-200a-3p, miR-200a-5p, miR-205-5p, miR-210-3p, miR-21-5p, miR-221-5p, miR-224-5p, miR-301a-3p, miR-31-5p, miR-34b-3p, miR-409-5p, miR-411-5p, miR-708-5p, miR-7-1-3p) specific miRNAs, as a combination of miRNAs, has more accurate predicted value in distinguishing cancer



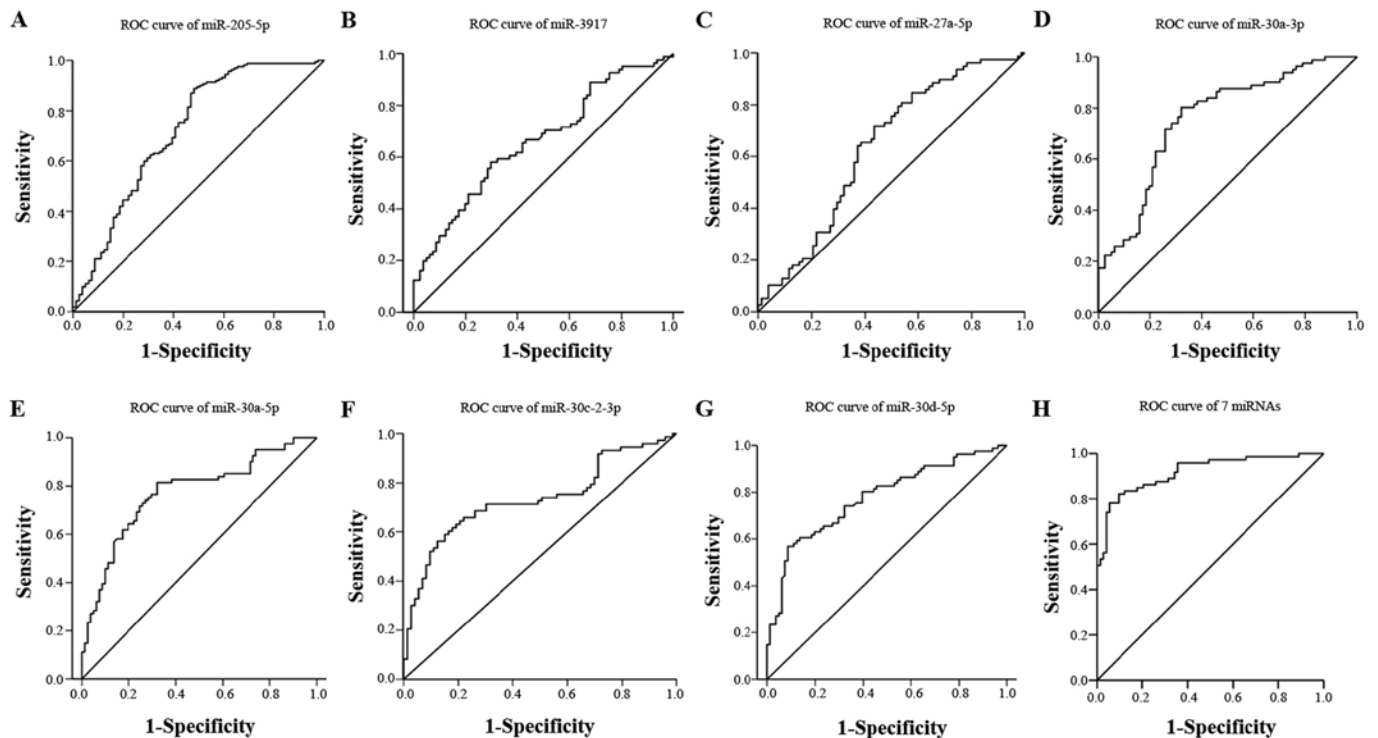


Figure 9. ROC curve analysis of signature miRNAs.

Table IV. ROC curve analysis of signature miRNAs.

miRNAs	Area	SE	P-value	95% CI
miR-205-5p	0.728	0.040	<0.001	0.650-0.806
miR-3917	0.661	0.042	<0.001	0.578-0.744
miR-27a-5p	0.637	0.045	0.003	0.549-0.725
miR-30a-3p	0.758	0.038	<0.001	0.683-0.832
miR-30a-5p	0.772	0.037	<0.001	0.699-0.845
miR-30c-2-3p	0.734	0.042	<0.001	0.650-0.817
miR-30d-5p	0.776	0.037	<0.001	0.704-0.848
Total 7 miRNAs	0.919	0.023	<0.001	0.874-0.963

cases from control. Pathway analysis results was indicative of dysregulated miRNAs related to Environmental Information Processing and Human Diseases, involved in non-small cell lung cancer. Studies have reported that these miRNAs are involved in the regulation of the development of lung cancer. For example, overexpression of miR-30a in lung adenocarcinoma A549 cells inhibited cell migration and invasion, which was partially through decrease of EYA2 expression (28). Chen *et al* (29) reported that miR-30d-5p was downregulated in lung cancer tissues, which significantly inhibited growth, cell cycle distribution, and motility of lung cancer cells. miR-30d-5p inhibited tumor cell proliferation and motility by directly targeting CCNE2 in NSCLC. miR-21 were reported to target MMP9 (30) responsible for cell proliferation and migration. Overall, these finding suggested that alterations of these miRNAs represent meaningful risk factors in lung cancer.

To validate the miRNA expression results from microarray and further confirm the relationship between differentially expressed miRNAs and lung cancer, we selected 7 miRNAs (miR-205-5p, miR-3917, miR-27a-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p and miR-30d-5p) which were found in non-small cell lung cancer pathway and analyzed their actual expression levels in 81 pairs of diagnosed lung tumor and adjacent non-tumor tissues by qRT-PCR. The results indicated that miR-205-5p and miR-3917 were upregulated and miR-27a-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p, miR-30d-5p were downregulated, which was highly consistent with the results of microarrays and TCGA analysis. miR-205-5p and miR-3917 has previously been described as upregulated in many types of cancer and tissues, including lung cancer, nasopharyngeal carcinoma and gingival tissues (31-33). miR-30a-3p and miR-30c-2-3p has been reported as downregulated in hepatocellular carcinoma and clear cell renal cell carcinomas (34,35). Then, we used conditional logistic regression analysis to further find the association between these signature miRNAs and lung cancer. The results showed that the increased risk of lung cancer was significantly associated with increased expression of miR-205-5p and miR-3917 and reduced expression of miR-27a-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p, and miR-30d-5p, respectively. ROC curve analysis demonstrated that miR-205-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p and miR-30d-5p could be considerable biomarkers for early diagnosis of lung cancer, also ROC analysis of all 7 miRNAs have more effective diagnosis for lung cancer. The results suggested that miR-205, miR-3917, miR-27a-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p, and miR-30d-5 may play an important role during lung cancer development and may be effective diagnostic biomarkers for lung cancer. Besides miR-27a, the function of the other six miRNAs in cancer has rarely been

reported. It is reported that miR-30a-5p promoted an epithelial phenotype and suppressed invasion by specifically targeting  $\beta 3$  integrin subunit to subsequently interdict the  $\beta 3$  integrin/Erk/Ets-1 network in breast cancer cells (36). miR-30a-5p might function as a tumor suppressing miRNA by directly targeting MTDH in hepatocellular carcinoma (37). Based on the above, the roles of these signature miRNAs in lung cancer still need further investigation in the future.

In conclusion, 28 miRNAs associated with lung cancer were identified by microarray, TCGA and miRNA-mRNA network analysis. After population analysis, miRNAs in miR-205, miR-3917, miR-27a-5p, miR-30a-3p, miR-30a-5p, miR-30c-2-3p, and miR-30d-5 were found to impact the risk of lung cancer and could be considerable biomarkers for early diagnosis of lung cancer. The results suggested a set of signature miRNAs may play a more important role during lung cancer development and promise biomarkers for the early screening of high-risk population. Integrating analysis including gene expression profile, TCGA database and population validation are effective methods to screen cancer related miRNAs. Future studies will investigate the roles of these miRNAs during the development of lung cancer.

## Acknowledgements

The present study was financially supported by the National Natural Science Foundation of China (81472939, 81673132, and 81182618), the Liu Da Ren Cai Gao Feng Project of Jiangsu Province (no. 2013-WSW-053) and the Fundamental Research Funds for the Central Universities.

## References

- Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J and Jemal A: Global cancer statistics, 2012. *CA Cancer J Clin* 65: 87-108, 2015.
- Chen W, Zheng R, Baade PD, Zhang S, Zeng H, Bray F, Jemal A, Yu XQ and He J: Cancer statistics in China, 2015. *CA Cancer J Clin* 66: 115-132, 2016.
- Vrijens K, Bollati V and Nawrot TS: MicroRNAs as potential signatures of environmental exposure or effect: A systematic review. *Environ Health Perspect* 123: 399-411, 2015.
- Zhang H, Li W, Nan F, Ren F, Wang H, Xu Y and Zhang F: MicroRNA expression profile of colon cancer stem-like cells in HT29 adenocarcinoma cell line. *Biochem Biophys Res Commun* 404: 273-278, 2011.
- Xia S, Huang CC, Le M, Dittmar R, Du M, Yuan T, Guo Y, Wang Y, Wang X, Tsai S, *et al*: Genomic variations in plasma cell free DNA differentiate early stage lung cancers from normal controls. *Lung Cancer* 90: 78-84, 2015.
- Pérez-Ramírez C, Cañadas-Garre M, Jiménez-Varo E, Faus-Dáder MJ and Calleja-Hernández MA: MET: A new promising biomarker in non-small-cell lung carcinoma. *Pharmacogenomics* 16: 631-647, 2015.
- Fan L, Qi H, Teng J, Su B, Chen H, Wang C and Xia Q: Identification of serum miRNAs by nano-quantum dots microarray as diagnostic biomarkers for early detection of non-small cell lung cancer. *Tumour Biol* 37: 7777-7784, 2016.
- Geng Q, Fan T, Zhang B, Wang W, Xu Y and Hu H: Five microRNAs in plasma as novel biomarkers for screening of early-stage non-small cell lung cancer. *Respir Res* 15: 149, 2014.
- Inamura K and Ishikawa Y: MicroRNA in lung cancer: Novel biomarkers and potential tools for treatment. *J Clin Med* 5: 5, 2016.
- Kim JO, Gazala S, Razzak R, Guo L, Ghosh S, Roa WH and Bédard EL: Non-small cell lung cancer detection using microRNA expression profiling of bronchoalveolar lavage fluid and sputum. *Anticancer Res* 35: 1873-1880, 2015.
- Wang Y, Zhang X, Liu L, Li H, Yu J, Wang C and Ren X: Clinical implication of microRNA for lung cancer. *Cancer Biother Radiopharm* 28: 261-267, 2013.
- Solomides CC, Evans BJ, Navenot JM, Vadigepalli R, Peiper SC and Wang ZX: MicroRNA profiling in lung cancer reveals new molecular markers for diagnosis. *Acta Cytol* 56: 645-654, 2012.
- Singh DK, Bose S and Kumar S: Role of microRNA in regulating cell signaling pathways, cell cycle, and apoptosis in non-small cell lung cancer. *Curr Mol Med* 16: 474-486, 2016.
- Razzak R, Bédard EL, Kim JO, Gazala S, Guo L, Ghosh S, Joy A, Nijjar T, Wong E and Roa WH: MicroRNA expression profiling of sputum for the detection of early and locally advanced non-small-cell lung cancer: A prospective case-control study. *Curr Oncol* 23: e86-e94, 2016.
- Pu Q, Huang Y, Lu Y, Peng Y, Zhang J, Feng G, Wang C, Liu L and Dai Y: Tissue-specific and plasma microRNA profiles could be promising biomarkers of histological classification and TNM stage in non-small cell lung cancer. *Thorac Cancer* 7: 348-354, 2016.
- Krutakova M, Sarlinova M, Matakova T, Dzian A, Hamzik J, Pec M, Javorkova S and Halasova E: The role of dysregulated microRNA expression in lung cancer. *Adv Exp Med Biol* 911: 1-8, 2016.
- Li JH, Liu S, Zhou H, Qu LH and Yang JH: starBase v2.0: Decoding miRNA-ceRNA, miRNA-ncRNA and protein-RNA interaction networks from large-scale CLIP-Seq data. *Nucleic Acids Res* 42: D92-D97, 2014.
- Hsu SD, Tseng YT, Shrestha S, Lin YL, Khaleel A, Chou CH, Chu CF, Huang HY, Lin CM, Ho SY, *et al*: miRTarBase update 2014: An information resource for experimentally validated miRNA-target interactions. *Nucleic Acids Res* 42: D78-D85, 2014.
- Li CY, Liang GY, Yao WZ, Sui J, Shen X, Zhang YQ, Peng H, Hong WW, Ye YC, Zhang ZY, *et al*: Integrated analysis of long non-coding RNA competing interactions reveals the potential role in progression of human gastric cancer. *Int J Oncol* 48: 1965-1976, 2016.
- Yang YL, Xu LP, Zhuo FL and Wang TY: Prognostic value of microRNA-10b overexpression in peripheral blood mononuclear cells of non-small-cell lung cancer patients. *Tumour Biol* 36: 7069-7075, 2015.
- Xu C, Zheng Y, Lian D, Ye S, Yang J and Zeng Z: Analysis of microRNA expression profile identifies novel biomarkers for non-small cell lung cancer. *Tumori* 101: 104-110, 2015.
- Wiwantkit V: MicroRNA assays for diagnosis lung cancer biopsy. *J Thorac Oncol* 10: e52-53, 2015.
- Wang P, Yang D, Zhang H, Wei X, Ma T, Cheng Z, Hong Q, Hu J, Zhuo H, Song Y, *et al*: Early detection of lung cancer in serum by a panel of microRNA biomarkers. *Clin Lung Cancer* 16: 313-319 e1, 2015.
- Chen X, Hu Z, Wang W, Ba Y, Ma L, Zhang C, Wang C, Ren Z, Zhao Y, Wu S, *et al*: Identification of ten serum microRNAs from a genome-wide serum microRNA expression profile as novel noninvasive biomarkers for non-small cell lung cancer diagnosis. *Int J Cancer* 130: 1620-1628, 2012.
- Yin QW, Sun XF, Yang GT, Li XB, Wu MS and Zhao J: Increased expression of microRNA-150 is associated with poor prognosis in non-small cell lung cancer. *Int J Clin Exp Pathol* 8: 842-846, 2015.
- Wu C, Cao Y, He Z, He J, Hu C, Duan H and Jiang J: Serum levels of miR-19b and miR-146a as prognostic biomarkers for non-small cell lung cancer. *Tohoku J Exp Med* 232: 85-95, 2014.
- Molina-Pinelo S, Pastor MD, Suarez R, Romero-Romero B, González De la Peña M, Salinas A, García-Carbonero R, De Miguel MJ, Rodríguez-Panadero F, Carnero A, *et al*: MicroRNA clusters: Dysregulation in lung adenocarcinoma and COPD. *Eur Respir J* 43: 1740-1749, 2014.
- Yuan Y, Zheng S, Li Q, Xiang X, Gao T, Ran P, Sun L, Huang Q, Xie F, Du J, *et al*: Overexpression of miR-30a in lung adenocarcinoma A549 cell line inhibits migration and invasion via targeting EYA2. *Acta Biochim Biophys Sin (Shanghai)* 48: 220-228, 2016.
- Chen D, Guo W, Qiu Z, Wang Q, Li Y, Liang L, Liu L, Huang S, Zhao Y and He X: MicroRNA-30d-5p inhibits tumour cell proliferation and motility by directly targeting CCNE2 in non-small cell lung cancer. *Cancer Lett* 362: 208-217, 2015.
- Hu J, Ni S, Cao Y, Zhang T, Wu T, Yin X, Lang Y and Lu H: The angiogenic effect of microRNA-21 targeting TIMP3 through the regulation of MMP2 and MMP9. *PLoS One* 11: e0149537, 2016.

31. Jiang M, Zhang P, Hu G, Xiao Z, Xu F, Zhong T, Huang F, Kuang H and Zhang W: Relative expressions of miR-205-5p, miR-205-3p, and miR-21 in tissues and serum of non-small cell lung cancer patients. *Mol Cell Biochem* 383: 67-75, 2013.
32. Wang F, Lu J, Peng X, Wang J, Liu X, Chen X, Jiang Y, Li X and Zhang B: Integrated analysis of microRNA regulatory network in nasopharyngeal carcinoma with deep sequencing. *J Exp Clin Cancer Res* 35: 17, 2016. doi: 10.1186/s13046-016-0292-4.
33. Stoecklin-Wasmer C, Guarnieri P, Celenti R, Demmer RT, Kebschull M and Papapanou PN: MicroRNAs and their target genes in gingival tissues. *J Dent Res* 91: 934-940, 2012.
34. Wang W, Lin H, Zhou L, Zhu Q, Gao S, Xie H, Liu Z, Xu Z, Wei J, Huang X, *et al*: MicroRNA-30a-3p inhibits tumor proliferation, invasiveness and metastasis and is downregulated in hepatocellular carcinoma. *Eur J Surg Oncol* 40: 1586-1594, 2014.
35. Moch H and Lukamowicz-Rajska M: miR-30c-2-3p and miR-30a-3p: New pieces of the jigsaw puzzle in HIF2 $\alpha$  regulation. *Cancer Discov* 4: 22-24, 2014.
36. Li W, Liu C, Zhao C, Zhai L and Lv S: Downregulation of  $\beta$ 3 integrin by miR-30a-5p modulates cell adhesion and invasion by interrupting Erk/Ets-1 network in triple-negative breast cancer. *Int J Oncol* 48: 1155-1164, 2016.
37. Li WF, Dai H, Ou Q, Zuo GQ and Liu CA: Overexpression of microRNA-30a-5p inhibits liver cancer cell proliferation and induces apoptosis by targeting MTDH/PTEN/AKT pathway. *Tumour Biol* 37: 5885-5895, 2016.