Original Paper

Three Circulating Long Non-Coding RNAs Act as Biomarkers for Predicting NSCLC

Qingfeng Tang[a,b]   Zhenhua Ni[b]   Zhuoan Cheng[c]   Jianhua Xu[c]   Hui Yu[d]   Peihao Yin[d]

[a]Department of Clinical Laboratory, Putuo Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, [b]Central Laboratory, Putuo Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, [c]Interventional Cancer Institute of Integrative Medicine, Putuo Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, [d]Department of General Surgery, Putuo Hospital, Shanghai University of Traditional Chinese Medicine, Shanghai, China

Abstract

Background/Aims: Circulating long non coding RNAs (lncRNAs) have emerged recently as major players in tumor biology and may be used for cancer diagnosis, prognosis, and as potential therapeutic targets. We explored circulating lncRNA as a predictor for the tumorigenesis of non-small-cell lung cancer (NSCLC). ***Methods:*** In this study, we applied a lncRNA microarray to screen for a potential biomarker for NSCLC, utilizing RT-PCR (ABI 7900HT). A multi-stage validation and risk score formula detection analysis was used. ***Results:*** We discovered that three lncRNAs (RP11-397D12.4, AC007403.1, and ERICH1-AS1) were up regulated in NSCLC, compared with cancer-free controls, with the merged area under the curve in the training and validation sets of 0.986 and 0.861. Furthermore, the positive predictive value and negative predictive value of the three merged factors were 0.72 and 0.87. We confirmed stable detection of the three lncRNAs by three cycles of freezing and thawing. ***Conclusions:*** RP11-397D12.4, AC007403.1, and ERICH1-AS1 may be potential biomarkers for predicting the tumorigenesis of NSCLC in the future.

Copyright © 2015 S. Karger AG, Basel

Introduction

Lung cancer is the leading cause of cancer-related mortality, with 1.4 million deaths reported annually worldwide [1]. Of these, non-small cell lung cancer (NSCLC) represents approximately 85% [2, 3] and has a dismal prognosis, with an overall 5-year survival rate of only 15% [4]. Therefore, identifying accurate predictive biomarkers is of great value in further understanding the biology of NSCLC and in developing novel therapeutic strategies. Thus, promising noninvasive biomarkers to identify high-risk individuals, in combination with diagnostic technology, is urgently needed.

Q. Tang, Z. Ni and Z. Cheng contributed equally to this work.

Peihao Yin

Department of General Surgery, Putuo Hospital, Shanghai University of Traditional Chinese Medicine, No.164 Lanxi Road, Shanghai 200062, (China)
Tel. +862185033768, Fax +862185033768, E-Mail yinpeihao1975@hotmail.com

KARGER 125 YEARS

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2015;37:1002-1009

DOI: 10.1159/000430226
Published online: September 23, 2015

© 2015 S. Karger AG, Basel
www.karger.com/cpb

1003

Tang et al.: Circulation LncRNA in NSCLC

The long non-coding RNAs (lncRNAs) have been reported as a biomarker for diagnosis, prognosis, and metastasis in multiple diseases [5-7]. For example, researchers have identified three circulation lncRNAs in predicting human hepatocellular carcinoma through high-throughput screening [8]. Although numerous studies have investigated small RNAs such as microRNAs (miRNAs) as potential biomarkers for the diagnosis of NSCLC [9-11], the diagnostic utility of circulating lncRNAs in NSCLC has never been explored. In this study, by using both Affymetrix (CA, USA) lncRNAs microarrays and reverse-transcription quantitative polymerase chain reaction (RT-qPCR) assays, we aimed to explore the genome-wide lncRNAs expression profile in plasma from NSCLC patients as compared with cancer-free controls, and thus sought to identify plasma lncRNAs that might serve as a novel biomarker panel in the diagnosis of NSCLC

Materials and Methods

Samples

The study totally enrolled 302 patients who had been pathological diagnosed as NSCLC in the Putuo Hospital (Shanghai, China) between 2009 and 2013. Informed consent for blood analysis was obtained prior to surgery; the study was approved by the ethics committee at the Putuo Hospital. All research was performed in compliance with government policies and the Helsinki Declaration. Experiments were undertaken with the understanding and written consent of each subject.

RNA extraction and RT-qPCR

Blood samples were collected from each donor were placed in the EDTA- anticoagulant tube. The plasma were separated by centrifugation at 800g for 10 min at room temperature, followed by a 15-min high-speed centrifugation at 10 000g at room temperature to completely remove cell debris. The supernatant plasma was recovered and stored at -80 °C until analysis. We extracted total RNA from 600μL plasma by Trizol reagent according to the protocol of manufactory (Invitrogen, CA, USA). During the progression of the RNA extraction, we added the cel-miR-39 into the RNA as the external normalization. The RNA has been qualified with a 260/280 value of 1.8-2.0. For the analysis of the expression levels of lncRNAs, both the internal reference (U6) and external normalization was applied for normalization. Reverse transcription was conducted by using random primer technology. RT-PCR was performed by using ABI 7900HT according to the protocol of manufactory in triplicate (ABI, CA, USA) as described previously [12].

Microarray analysis of lncRNAs

Total RNA from each sample was quantified by the NanoDrop ND-1000 (Thermo, CA, USA). For microarray analysis, Agilent Array platform was employed. The sample preparation and microarray hybridization were performed based on the manufacturer's standard protocols with minor modifications. Then, each sample was amplified and transcribed into fluorescent cRNA along the entire length of the transcripts without 3' bias utilizing a random priming method.

Risk score analysis

Risk score analysis was performed to evaluate the associations between the concentrations of the plasma lncRNA expression levels. The upper 95% reference interval of each lncRNA value in controls was set as the threshold to code the expression level of the corresponding lncRNA for each sample as 0 and 1 in the training set. A risk score function (RSF) to predict NSCLC was defined according to a linear combination of the expression level for each lncRNA. For example, the RSF for sample i using information from three lncRNAs was: $rsf_i = \sum 3j{-}1 Wj.s_{ij}$. In the above equation, s_{ij} is the risk score for lncRNA j on sample i, and W_j is the weight of the risk score of lncRNA j. We conducted the ROC analysis by using the total RSF value according to the case-control group in training set. We chose the 8.626 as the cutoff value because the value of sensitivity + specificity was considered to be maximal.

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2015;37:1002-1009

DOI: 10.1159/000430226
Published online: September 23, 2015

© 2015 S. Karger AG, Basel
www.karger.com/cpb

1004

Tang et al.: Circulation LncRNA in NSCLC

Results

Microarray based detection of plasma lncRNAs

Human lncRNA Array (version 3.0, Agilent, CA, USA) was applied to detect the lncRNA derived from the plasma of 11 patients with NSCLC, and a cancer-free cohort consisting of nine individuals. The clinicopathologic relevance analysis (total of 507 patients) is summarized in Table 1. All 302 patients enrolled in this study were clinically and pathologically diagnosed with NSCLC. There were no significant differences in the distribution of age and sex between cancer patients and cancer-free controls.

A total of 803 lncRNA transcripts were specifically deregulated, including 601 up regulated lncRNA transcripts and 202 down regulated lncRNA transcripts; each $p < 0.05$) in patients with NSCLC compared with cancer-free controls (Fig. 1A). In order to screen the biomarker predication to NSCLC tumorigenesis, a 4/0.25 fold change was used as the cutoff. We obtained 32 up regulated lncRNAs and 10 down regulated lncRNAs. Next, filtering all 42 deregulated transcripts for high signal intensity (≥ 8) and at least fourfold deregulation yielded eight lncRNA candidates as listed (Fig. 1B). Eight candidate lncRNAs were selected to be validated in the 20 pairs of samples of the following training set.

Validation by training set and validation set

We further examined these differentially expressed lncRNAs by RT-qPCR in a training sample set, including 20 cases and 20 controls (including the same samples used in the microarray assays). As shown in Fig. 1B, this phase generated a panel of three lncRNAs (RP11-397D12.4, AC007403.1, and ERICH1-AS1) that were significantly up regulated in NSCLC samples.

To validate the accuracy and specificity of these three lncRNA as a NSCLC potential signature, we also examined their expression levels through a larger cohort of individual samples (232 cases and 135 controls). As shown in Fig. 2A, the expression of three lncRNAs in the plasma of NSCLC patients was significantly higher than in the control group. Throughout the multiphase testing and analysis phases, a profile of three lncRNAs might be considered to be the potential signature for NSCLC tumorigenesis.

The risk score formula was applied to assess the diagnostic value of the three lncRNAs profiling system. First, the risk score of each plasma sample in the training set was calculated, as the basis of their risk scores and a set cutoff value; plasma samples were then divided into a high-risk group, representing the possible NSCLC group, and a low-risk group, representing the predicted controls. At the optimal cutoff value (8.626), with the value of sensitivity + specificity considered to be maximal, the positive predictive value (PPV) and negative predictive value (NPV) were 95% and 100% in the training set, respectively. Similarly, when the same cutoff value was applied to calculate the risk score of samples in the larger validation sets, the PPV and NPV were 72% and 87%, respectively (Table 2).

The ROC curves analysis was also conducted to assess the diagnostic sensitivity and specificity of the three lncRNAs signature for NSCLC by using risk score functions (RSF).

Table 1. Clinicopathological features of surgical non-small-cell lung cancer (NSCLC) and cancer-free control samples. [a]Student t-test, [b]Chi-square test, [c]squamous cell carcinoma, [d]adenocarcinoma

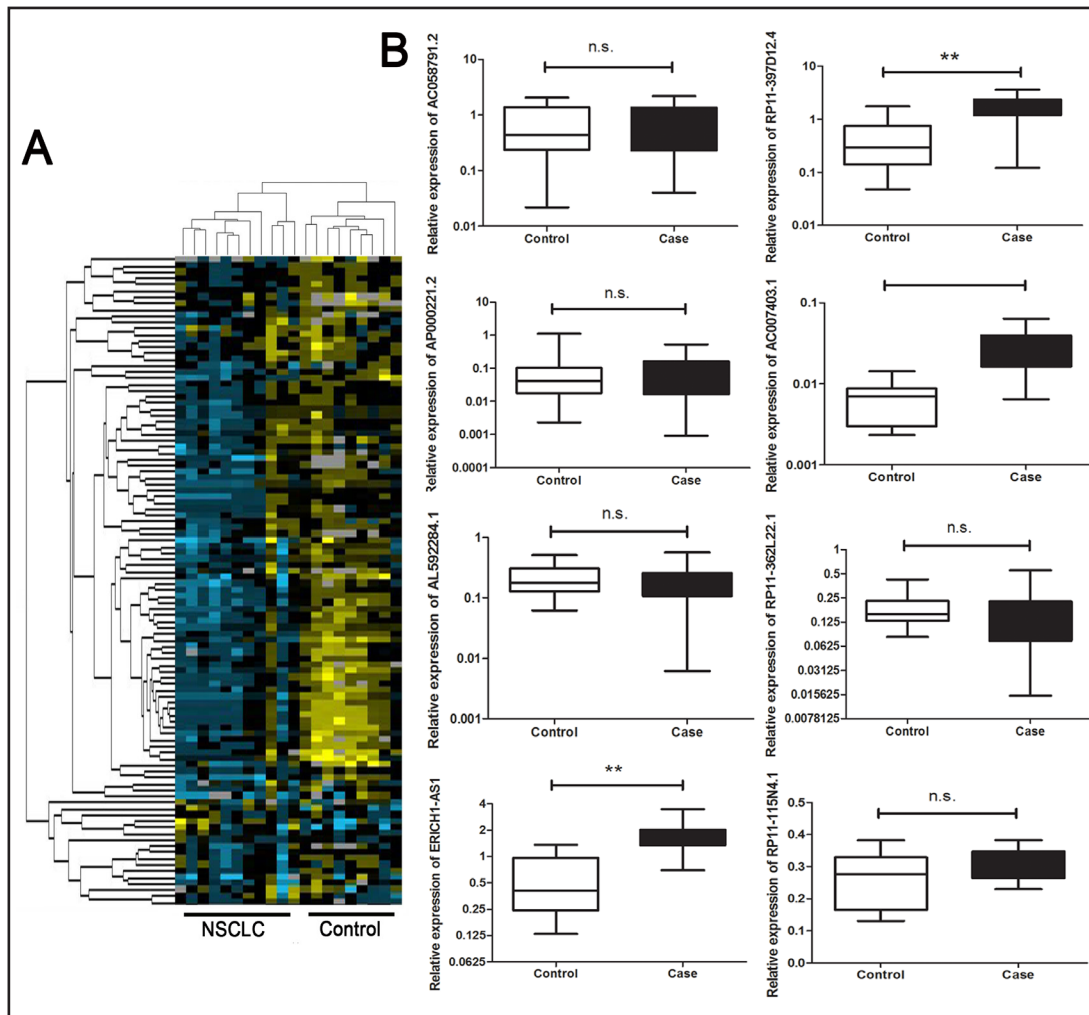| N	NSCLC 302	Control 205	P value
Age Mean (SE) year	59.11(0.66)	59.18(0.54)	0.21[a]
Sex (male/female)	262/40	175/30	0.66[b]
Smoking history			
No[c]			
Yes[d]			
Histology subtype			
SCC	147		
ADC	112		
Other	43		
Tumor Size(cm)			
≤5cm	197		
>5cm	105		
Tumor Number			
Solitary	201		
Multiple	101		
Tumor Capsular			
Incomplete	52		
Complete	250		
TNM stage			
I-II	230		
III-IV	72		

Fig. 1. Expression of the candidate lncRNAs based on the high throughput detection. (A) Hierarchical cluster analysis of different expressed lncRNAs detected by high throughput microarray. (B) Total 20 paired plasma from NSCLC patients20 cancer-free controls were used in validating expression of eight candidate lncRNAs including AC058791.2, RP11-397D12.4, AL592284.1, RP11-362L22.1, AP000221.2, AC007403.1, ERICH1-AS1, RP11-115N4.1 during RT-qPCR analysis. Data was presented as mean ± SEM. Data was analyzed with student t test. ** indicated p < 0.01.

Single lncRNA and merged factors were analyzed. As presented in Fig. 2 (B and D), the areas under the curve (AUC) of RP11-397D12.4, AC007403.1, and ERICH1-AS1 and their merged factors were 0.900, 0.974, 0.875, and 0.996, respectively in the training set. In the validation set, the AUCs were 0.770, 0.868, 0.849, and 0.942, respectively (Fig. 2 C and E).

Double-blind test for validating the diagnostic capability
We tested another 100 plasma samples (50 NSCLC patients and 50 cancer-free controls) in a double-blinded fashion to validate the accuracy of the three plasma lncRNA biomarkers in diagnosing NSCLC. A clear separation of NSCLC cases from controls was observed, with an accuracy rate of 89%.

Stability expression of RP11-397D12.4, AC007403.1 and ERICH1-AS1 in human plasma
The expression of the three lncRNAs was detected from three healthy controls at room temperature for 12 and 24 h, or subjecting it to up to three cycles of freezing and thawing.
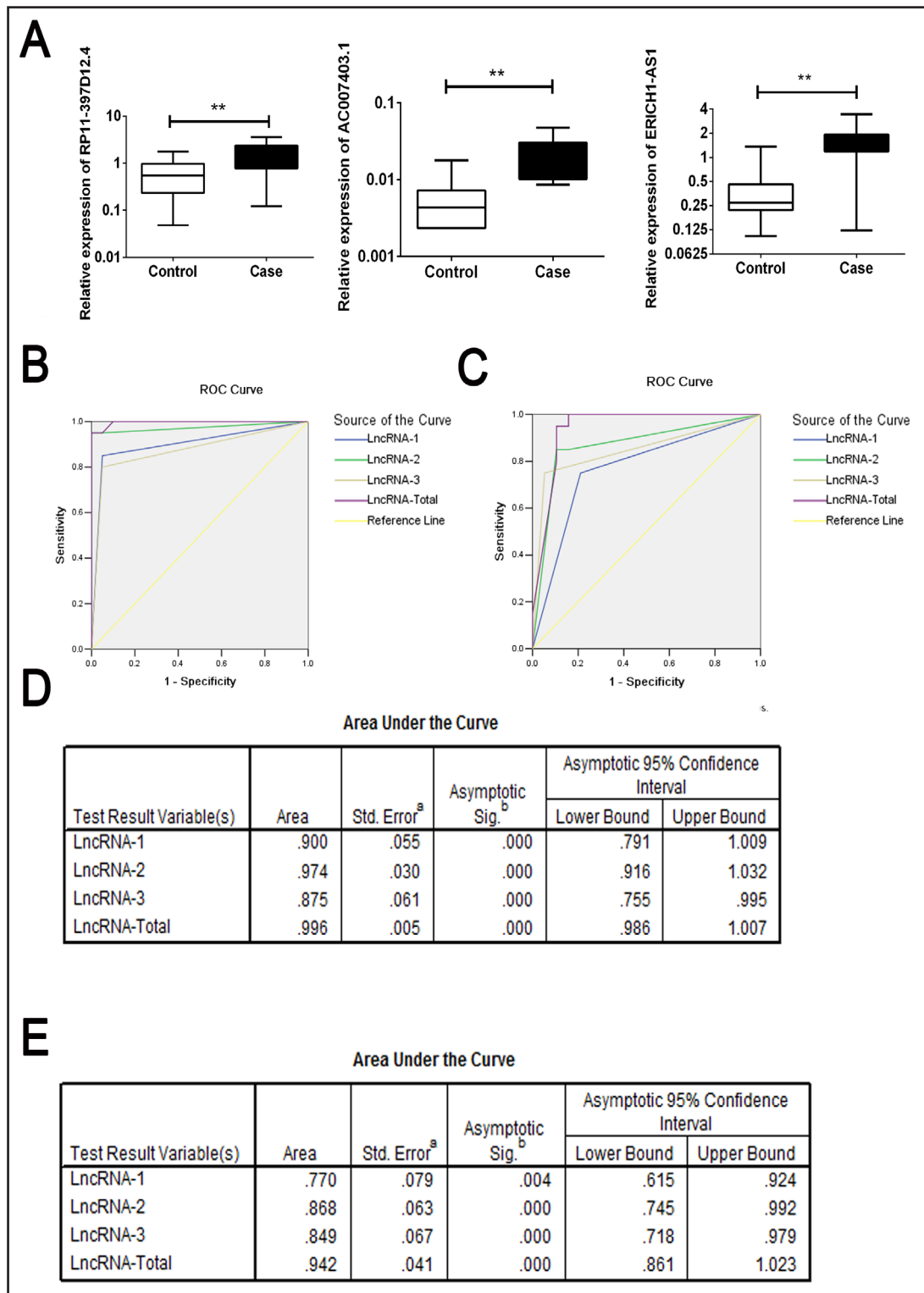
KARGER

Fig. 2. ROC curve analysis for predicting NSCLC. (A) Total 232 plasma from NSCLC patients and 135 cancer-free controls were used in RT-qPCR analysis. Data was presented as mean ± SEM and was analysis with student t test. ** indicated p < 0.01. (B, D) ROC curve analysis was conducted for discrimination between NSCLC and cancer-free controls by the three-lncRNA profile to separate 20 pairs in the training set. (C, E) ROC curve analysis was conducted for discrimination between NSCLC and cancer-free controls by the three-lncRNA profile in the validation set.

D

Area Under the Curve

Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
LncRNA-1	.900	.055	.000	.791	1.009
LncRNA-2	.974	.030	.000	.916	1.032
LncRNA-3	.875	.061	.000	.755	.995
LncRNA-Total	.996	.005	.000	.986	1.007

E

Area Under the Curve

Test Result Variable(s)	Area	Std. Error[a]	Asymptotic Sig.[b]	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
LncRNA-1	.770	.079	.004	.615	.924
LncRNA-2	.868	.063	.000	.745	.992
LncRNA-3	.849	.067	.000	.718	.979
LncRNA-Total	.942	.041	.000	.861	1.023

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2015;37:1002-1009

DOI: 10.1159/000430226
Published online: September 23, 2015

© 2015 S. Karger AG, Basel
www.karger.com/cpb

1007

Tang et al.: Circulation LncRNA in NSCLC

Table 2. Risk score analysis of in NSCLC and cancer-free control plasma samples. [a]PPV, positive predictive value, [b]NPV, negative predictive value

Score	0-8.626	8.626-19.22	PPV [a]	NPV [b]
Training set			1.00	0.95
NSCLC	1	19		
Control	20	0		
Validation set			0.72	0.87
NSCLC	65	167		
Control	117	18		
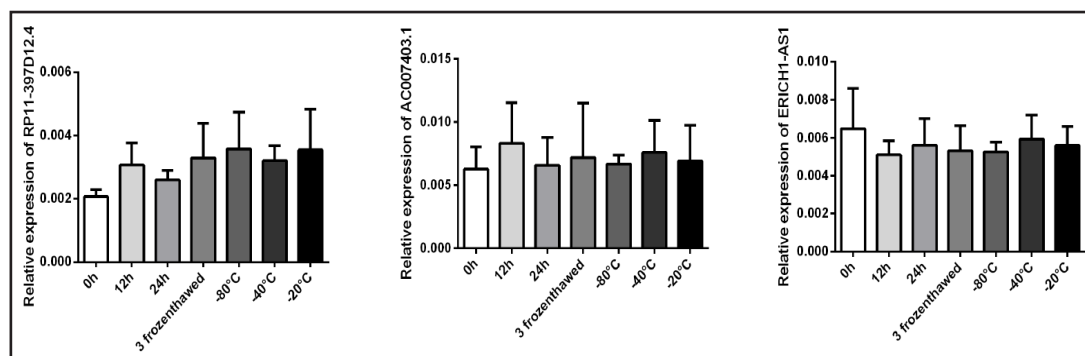
Fig. 3. Stability detection of the three lncRNAs in human plasma. RT-qPCR was applied for detecting the expression level of the three lncRNAs. Human plasmas obtained from three healthy controls were incubated at room temperature for 12h, 24h, subjecting it to up to 3 cycles of freezing and thawing or stored at -80°C for about 7 days or at -40°C, -20°C for 48h. Data was presented as mean ± SEM. No significant difference was observed in each group.

We also detected the expression level of the three lncRNAs in three volunteers, in storage conditions of –80° C for 7 days. The expression of lncRNAs stored at –40° C and -20° C for 48 h was also detected. The entire process had minimal effects on the concentrations of the three lncRNAs, demonstrating that they are sufficiently stable in human plasma (Fig. 3).

Discussion

In the present study, we demonstrated that the expression level of RP11-397D12.4, AC007403.1, and ERICH1-AS1 was significantly different among NSCLC patients and healthy volunteers, and that their expression profiles can be used to differentiate NSCLC from healthy controls. In addition, the diagnostic performance of the combination panel of these three lncRNA (AUC, 0.942) was significantly higher than that of CEA [13, 14].

By using detection analyses such as RNA sequencing and microarray, altered expression of lncRNAs has been reported in various malignancies. Multiple deregulated lncRNAs have also been implicated in physiologic and pathologic processes of lung cancer [15-18]. Based on deregulated lncRNAs expression profiling and their association with the biologic and clinical properties of lung cancer, specific lncRNAs could be utilized to distinguish benign and malignant lesions [19]. Given that lncRNAs are secreted from tumor cells into the circulatory system, they could be developed into a novel, reliable blood-based fingerprinting system for lung cancer diagnosis [20]. Herein, we performed a case-control study through the high-throughput lncRNA microarray, and discovered three novel lncRNA, RP11-397D12.4, AC007403.1, and ERICH1-AS1, acting as predictive biomarker for NSCLC. The risk score analysis, including multistage validation, was employed to evaluate the association between NSCLC and the lncRNAs expression levels. Additionally, deep-sequencing based characterization of exosomal RNA from human plasma revealed that more than 3% of total exosomal RNA is represented by lncRNAs [21-23]. The exosome might be the possible vector of the lncRNAs in hematogenous dissemination. The mechanisms accounting for the

Cellular Physiology and Biochemistry

Cell Physiol Biochem 2015;37:1002-1009

DOI: 10.1159/000430226
Published online: September 23, 2015

© 2015 S. Karger AG, Basel
www.karger.com/cpb

1008

Tang et al.: Circulation LncRNA in NSCLC

stability of plasma lncRNAs are not well understood; they may be protected by exosome encapsulation, such as plasma microRNAs [24].

In conclusion, we identified three lncRNAs, RP11-397D12.4, AC007403.1, and ERICH1-AS1, as potential fingerprints for tumorigenesis prediction. We propose that this panel of lncRNAs might be utilized to develop noninvasive screening tools for NSCLC. More in-depth studies are required to confirm the diagnostic values of the three lncRNAs in discriminating NSCLC from benign pulmonary disease.

Acknowledgements

Disclosure Statement

The authors declare that they have no financial conflict of interest.

References

1   Smith RA, Manassaram-Baptiste D, Brooks D, Doroshenk M, Fedewa S, Saslow D, Brawley OW, Wender R: Cancer screening in the United States, 2015: a review of current American cancer society guidelines and current issues in cancer screening. CA Cancer J Clin 2015;65:30-54.

2   Ettinger DS, Wood DE, Akerley W, Bazhenova LA, Borghaei H, Camidge DR, Cheney RT, Chirieac LR, D'Amico TA, Demmy TL, Dilling TJ, Dobelbower MC, Govindan R, Grannis FW, Jr., Horn L, Jahan TM, Komaki R, Krug LM, Lackner RP, Lanuti M, Lilenbaum R, Lin J, Loo BW, Jr., Martins R, Otterson GA, Patel JD, Pisters KM, Reckamp K, Riely GJ, Rohren E, Schild SE, Shapiro TA, Swanson SJ, Tauer K, Yang SC, Gregory K, Hughes M: Non-small cell lung cancer, version 6.2015. J Natl Compr Canc Netw 2015;13:515-524.

3   Jacobs J, Zwaenepoel K, Rolfo C, Van den Bossche J, Deben C, Silence K, Hermans C, Smits E, Van Schil P, Lardon F, Deschoolmeester V, Pauwels P: Unlocking the potential of CD70 as a novel immunotherapeutic target for non-small cell lung cancer. Oncotarget 2015;6:13462-13475.

4   Dempke WC: Targeted Therapy for NSCLC-A Double-edged Sword? Anticancer Res 2015;35:2503-2512.

5   Wang Y, Gao S, Liu G, Jia R, Fan D, Feng X: Microarray expression profile analysis of long non-coding RNAs in human gastric cardiac adenocarcinoma. Cell Physiol Biochem 2014;33:1225-1238.

6   Tong YS, Wang XW, Zhou XL, Liu ZH, Yang TX, Shi WH, Xie HW, Lv J, Wu QQ, Cao XF: Identification of the long non-coding RNA POU3F3 in plasma as a novel biomarker for diagnosis of esophageal squamous cell carcinoma. Mol Cancer 2015;14:3.

7   Ronnau CG, Verhaegh GW, Luna-Velez MV, Schalken JA: Noncoding RNAs as novel biomarkers in prostate cancer. Biomed Res Int 2014;2014:591703.

8   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;6:4505-4515.

9   Zhang L, Qian J, Qiang Y, Huang H, Wang C, Li D, Xu B: Down-regulation of miR-4500 promoted non-small cell lung cancer growth. Cell Physiol Biochem 2014;34:1166-1174.

10   Yu H, Jiang L, Sun C, Li Guo L, Lin M, Huang J, Zhu L: Decreased circulating miR-375: a potential biomarker for patients with non-small-cell lung cancer. Gene 2014;534:60-65.

11   Markou A, Sourvinou I, Vorkas PA, Yousef GM, Lianidou E: Clinical evaluation of microRNA expression profiling in non small cell lung cancer. Lung Cancer 2013;81:388-396.

12   Li J, Chen Z, Tian L, Zhou C, He MY, Gao Y, Wang S, Zhou F, Shi S, Feng X, Sun N, Liu Z, Skogerboe G, Dong J, Yao R, Zhao Y, Sun J, Zhang B, Yu Y, Shi X, Luo M, Shao K, Li N, Qiu B, Tan F, Chen R, He J: LncRNA profile study reveals a three-lncRNA signature associated with the survival of patients with oesophageal squamous cell carcinoma. Gut 2014;63:1700-1710.

13 Jiang AG, Chen HL, Lu HY: The relationship between glasgow prognostic score and serum tumor markers in patients with advanced non-small cell lung cancer. BMC Cancer 2015;15:386.

14 Romero-Ventosa EY, Blanco-Prieto S, Gonzalez-Pineiro AL, Rodriguez-Berrocal FJ, Pineiro-Corrales G, Paez de la Cadena M: Pretreatment levels of the serum biomarkers CEA, CYFRA 21-1, SCC and the soluble EGFR and its ligands EGF, TGF-alpha, HB-EGF in the prediction of outcome in erlotinib treated non-small-cell lung cancer patients. Springerplus 2015;4:171.

15 Yu H, Xu Q, Liu F, Ye X, Wang J, Meng X: Identification and validation of long noncoding RNA biomarkers in human non-small-cell lung carcinomas. J Thorac Oncol 2015;10:645-654.

16 Sun M, Liu XH, Lu KH, Nie FQ, Xia R, Kong R, Yang JS, Xu TP, Liu YW, Zou YF, Lu BB, Yin R, Zhang EB, Xu L, De W, Wang ZX: EZH2-mediated epigenetic suppression of long noncoding RNA SPRY4-IT1 promotes NSCLC cell proliferation and metastasis by affecting the epithelial-mesenchymal transition. Cell Death Dis 2014;5:e1298.

17 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;85:110-115.

18 Whiteside EJ, Seim I, Pauli JP, O'Keeffe AJ, Thomas PB, Carter SL, Walpole CM, Fung JN, Josh P, Herington AC, Chopin LK: Identification of a long non-coding RNA gene, growth hormone secretagogue receptor opposite strand, which stimulates cell migration in non-small cell lung cancer cell lines. Int J Oncol 2013;43:566-574.

19 Yang Y, Li H, Hou S, Hu B, Liu J, Wang J: The noncoding RNA expression profile and the effect of lncRNA AK126698 on cisplatin resistance in non-small-cell lung cancer cell. PLoS One 2013;8:e65309.

20 Tang J, Zhuo H, Zhang X, Jiang R, Ji J, Deng L, Qian X, Zhang F, Sun B: A novel biomarker Linc00974 interacting with KRT19 promotes proliferation and metastasis in hepatocellular carcinoma. Cell Death Dis 2014;5:e1549.

21 Pefanis E, Wang J, Rothschild G, Lim J, Kazadi D, Sun J, Federation A, Chao J, Elliott O, Liu ZP, Economides AN, Bradner JE, Rabadan R, Basu U: RNA exosome-regulated long non-coding RNA transcription controls super-enhancer activity. Cell 2015;161:774-789.

22 Shah S, Wittmann S, Kilchert C, Vasiljeva L: lncRNA recruits RNAi and the exosome to dynamically regulate pho1 expression in response to phosphate levels in fission yeast. Genes Dev 2014;28:231-244.

23 Kim NH, Choi SH, Kim CH, Lee CH, Lee TR, Lee AY: Reduced MiR-675 in exosome in H19 RNA-related melanogenesis via MITF as a direct target. J Invest Dermatol 2014;134:1075-1082.

24 Senfter D, Holzner S, Kalipciyan M, Staribacher A, Walzl A, Huttary N, Krieger S, Brenner S, Jager W, Krupitza G, Dolznig H, Mader RM: Loss of miR-200 family in 5-fluorouracil resistant colon cancer drives lymphendothelial invasiveness *in vitro*. Hum Mol Genet 2015;24:3689-3698.