Circulating plasma microRNAs as potential biomarkers of non-small cell lung cancer obtained by high-throughput real-time PCR profiling

Yanqin Niu¹*, Mingyang Su¹,²*, Yike Wu¹, Liwu Fu³, Kang Kang⁴, Qing Li¹, Li Li¹,
Gang Hui⁵, Furong Li⁶# and Deming Gou¹#

¹Shenzhen Key Laboratory of Microbial Genetic Engineering, College of Life Sciences and Oceanography, Carson International Cancer Center, Shenzhen University, Shenzhen, Guangdong, 518060, China;
²Key Laboratory of Optoelectronic Devices and Systems of Ministry of Education and Guangdong Province, College of Optoelectronic Engineering Shenzhen University, Shenzhen, Guangdong, 518060, China;
³Collaborative Innovation Center for Cancer Medicine, State Key Laboratory of Oncology in South China, Sun Yat-sen University Cancer Center, Guangzhou 510060, China.
⁴Department of Biochemistry and Molecular Biology, School of Basic Medical Sciences, Shenzhen University Health Sciences Center, Shenzhen, Guangdong, 518060, China.
⁵Department of Thoracic Surgery, Peking University Shenzhen Hospital, Shenzhen, Guangdong, 518036, China.
⁶Translational Medicine Collaborative Innovation Center, The Second Clinical Medical College (Shenzhen People’s Hospital), Jinan University, Shenzhen 510632, China.

Running title: Circulating microRNA, biomarker, NSCLC, real-time PCR
Keywords: NSCLC, circulating microRNA, biomarker, S-Poly(T) Plus, real-time PCR

Grants

This work was supported by National Natural Science Foundation of China (No.81600039, 91739109, 81700054, 81570046 and 31571199), Shenzhen Municipal Basic Research Program (JCYJ20150729104027220, JCYJ20170818144127727 and CXZZ20130515092016300), Shenzhen University start-up funding (2018043) and Shenzhen University Interdisciplinary Innovation Team Project (000003).

# To whom correspondence should be addressed: Deming Gou, College of Life Sciences and Oceanography, Shenzhen University, Nanhai Ave 3688, Shenzhen, Guangdong, 518060, China. Tel: 86-755-26527848; Fax: 86-755-26534274; E-mail: dmgou@szu.edu.cn.

Furong Li, Translational Medicine Collaborative Innovation Center, The Second Clinical Medical College (Shenzhen People’s Hospital), Jinan University, Shenzhen, 510632, China. Tel: 86-755-25533018; Fax: 86-755-25533018; E-mail: frli62@163.com.

* These authors contributed equally to this work.

Conflict of Interest Statement

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

Background

Because of limited stability and sensitivity, circulating miRNAs as noninvasive biomarkers has not so far been used for early diagnosis and prognosis of non-small cell lung cancer (NSCLC) in clinic. Therefore, it is still imperative to find more reliable biomarker(s).

Methods

We performed one of most sensitive RT-qPCR assays, S-Poly(T) Plus to selected differently expressed miRNAs from genome-wide miRNA profiling. miRNA candidates were validated through a three-phase selection and two-validation process with 437 NSCLC cases and 415 controls.

Results

A unique set of 7 and 9 miRNAs differed significantly in adenocarcinoma (ADC) and squamous cell carcinoma (SCC) samples compared with those in controls, of which, there were 5 universal biomarkers for NSCLC (ADC or SCC). Ten out of eleven miRNAs could discriminate early stage (Stage I) of NSCLC from healthy individuals. Risk score was obtained from the validation set-1 and was tested using operating characteristic (ROC) curves with a high area under ROC curve of 0.89 in ADC and 0.96 in SCC. Ultimately, potential biomarkers and the risk score were verified by the validation set-2 with a sensitivity of 94% and a specificity of 91.6% in ADC, and a sensitivity of 98.5% and a specificity of 51.5% in SCC, respectively.

Conclusion
Taken together, 7 miRNAs and 9 miRNAs may provide noninvasive biomarkers for diagnosis and prognosis in ADC and SCC, respectively.

Impact

Based on our sensitive and accurate method, we hope that these candidate miRNAs may have strong impact on the early lung cancer diagnosis.

Keywords: NSCLC, circulating microRNA, biomarker, S-Poly(T) Plus, real-time PCR
Introduction

Lung cancer is the leading cause of cancer deaths in the world and 80% patients have non-small cell lung cancer (NSCLC), which often diagnosed at later stages with around 15% of 5-year survival rates. If discovered at stage I, 60-80% of patients could be survived (1). Unfortunately, current approaches to cancer screening are invasive and it is difficult to detect cancer at very early stage. Subsequently, the discovery and application of molecular biomarkers has opened new sights for diagnosis and prediction of cancer. Recently, advances in miRNA biomarkers have generated some candidate markers of lung cancer with potential clinical values (1-2). Moreover, it was reported that miRNA expression analyses in plasma samples collected 1-2 years before the onset of disease, at the time of CT (computed tomography) detection and in disease-free smokers, resulted in the generation of miRNA signatures with strong predictive, diagnostic, and prognostic potential (3). However, many previously published studies tested either single miRNA as diagnostic marker or validated in rather small patient cohorts. In addition, most of studies lacked validation in different cohorts. Until today, these published results have not been applied for clinic purpose. Therefore, it is still imperative to find reliable biomarker(s) with stability and reproducibility for clinical applications, thus improving the diagnosis and prognosis of this disease.

MicroRNAs (miRNAs) are small noncoding RNA (~22 nucleotides), regulating gene expression at the post-transcriptional level by binding the 3’-untranslated region (3’-UTR) of target mRNAs (4). Physiologically, miRNAs have relevance to cell differentiation, proliferation and apoptosis and many diseases including cancer (5-7). One miRNA may regulate hundreds of mRNAs, as a result, miRNA expression patterns can be surprisingly informative (8), to an extent, may reflect the
differentiation state of the tumors (9). miRNAs were originally studied in tissue, but multiple studies have demonstrated that miRNAs are detectable and highly stable in the circulation (10). Given the above, circulating miRNAs show great potential as novel noninvasive biomarkers for cancers (11-13).

To determine the expression pattern of miRNAs, an accurate and inexpensive profiling approach is needed. Microarray or RNA sequence have been used for miRNA high-throughput profiling (14), but the sensitivity and dynamic range have been problematic and are therefore best suited as discovery tools but quantitative assay platforms (8). The quantitative real-time PCR (qRT-PCR) based miRNA expression assay might be one of most sensitive techniques for miRNA detection (15), of which, the unique design of S-Poly(T) Plus method showed increased sensitivity when compared to others (16, 17). In addition, the S-Poly(T) Plus also has the attractive properties of increased simplicity, high efficiency and low cost, allowing detection of expression pattern of more miRNAs across a large panel of samples.

In this article we aim to use genome-wide expression profile from large cohort of NSCLC patients to conduct an unbiased approach for identification of potentially prognostic circulating miRNAs. First, extensive literatures search was performed. The keyword “microRNA/miRNA” was combined with the search string “cancer”. A total of 486 miRNAs were collected from literatures based on searched results. Afterwards, these miRNAs were assessed with large-scale data sets from patients with NSCLC and healthy volunteers; altered miRNA subsets were subsequently quantified in samples from 288 patients with NSCLC and 266 healthy donors, and then potential biomarkers and risk score were revalidated in different set of plasma samples from 149 NSCLC patients and 149 healthy controls. Considering the accessibility and
stability of circulating miRNAs, 7 miRNAs and 9 miRNAs provide noninvasive biomarkers for diagnosis and prognosis in ADC and SCC, respectively.

Materials and methods

Plasma collection

To control the disease heterogeneity, we used patients with two common histologic subtypes of NSCLC, adenocarcinoma (ADC) and squamous cell carcinoma (SCC). Demographic and etiologic characteristics of ADC and SCC have been described (18, 19). All patients were first diagnosed and the blood was collected ahead of any clinic treatments. Whole blood samples were put into EDTA-containing tubes individually and then centrifuged at 3,000 × g for 10 min at 4°C. The supernatants were transferred into RNase-free tubes and stored at -80°C. Given that miRNAs also expressed in red blood cells and could have influence on the assessment, all haemolysis of blood samples was excluded from consideration. Plasma samples were collected and examined under blinded testing. Qualified plasma samples were prospectively collected from 266 healthy donors and 288 patients with non-small cell lung cancer (NSCLC) in Shenzhen People’s Hospital (Shenzhen, China), as well as from 149 patients with NSCLC and 149 healthy donors in Cancer Center of Guangzhou Medical University (Guangzhou, China). This study had been approved by the Institutional Ethics Committee at Shenzhen People’s Hospital and Cancer Center of Guangzhou Medical University, and informed written consents were obtained from all subjects.

RNA extraction
Total RNA was isolated using our developed serum/plasma (S/P) miRsol method as previously described (16). We added 0.1 pM spiked-in *Caenorhabditis elegans* cel-miR-54-5p to 100 μl of plasma for normalization.

**Polyadenylation, reverse transcription and Real-time PCR**

The polyadenylation, reverse transcription and real-time PCR procedure were conducted exactly according to S-Poly(T) Plus protocol (16). To profile all miRNAs effectively, every 7 out of 486 miRNAs were grouped together, and were polyadenylated and reverse transcribed (RT) simultaneously in one reaction. miRNAs with identical forward primers or with more than five base-pairings between the forward primer and RT primers should be avoided in a group. All sequences, primers and probes were listed in the supplemental table 1.

Real-time PCR was performed in 96-well plates by using ABI StepOnePlus thermal cycler as previously described (16). Each PCR reaction was carried out in duplicate. The miRNA expression level was normalized to spiked-in cel-miR-54. miRNAs with cycle threshold (Ct) value less than 35 in the panel were included in data analysis.

**miRNA profiling**

Excellent biomarkers should be qualified in the mixed and individual samples, and in different sets of samples. Therefore, a five-step test was designed to identify prognostic miRNAs for NSCLC, and the whole study flow chart was depicted as in figure 1. First, 554 plasma samples from Shenzhen People’s Hospital (Shenzhen, China) were enrolled into three cohorts, which were healthy control (N=266), NSCLC stage I (N=130) and stage II-IV (N=158), and afterwards pooled together respectively.
miRNA profiling assay was performed using the S-Poly(T) Plus approach with every single miRNA being detected from each pool. Comparing those in healthy cohort, miRNAs in group stage I or group stage II-IV differ by more than 4-fold changes on outcome were selected and secondly confirmed. In this step, all PCR products were detected by electrophoresis in 3.5% agarose gel, and miRNAs with nonspecific amplification were excluded; the second screening (further-screening) consisted of five cohorts: healthy control (N=266), ADC stage I (N=96), ADC stage II-IV (N=113), SCC stage I (N=34) and SCC stage II-IV (N=45). Comparing to those in healthy group, miRNAs from stage I or stage II-IV with 2-fold changes were further studied in ADC and SCC, respectively. Third, candidate miRNAs obtained from the further-screening were validated using small number of individual specimens randomly selected from Shenzhen People’s Hospital. The selection criterion of the training set was the same as that in further-screening. Ultimately, selected miRNAs were validated using each individual of 266 healthy donors and 288 patients with NSCLC from Shenzhen People’s Hospital, and then further validated in 149 patients with NSCLC and 149 healthy donors from Cancer Center of Guangzhou Medical University.

Statistics

Statistical analysis was submitted to the GraphPad Prism 5. It was unpaired two-tailed Student’s t test that used to compare the miRNA expressions in plasma, and paired two-tailed Student’s t test in tissue. Data were present as means ± SE (standard error), and variables with $p$ values less than 0.05 were retained. Bivariate logistic
regression analysis was performed using SPSS software, version 19.0 (SPSS Inc., Chicago, IL, USA) and the method was “enter”.

Results

Studied Participants and miRNAs

In this study, 486 miRNAs related to cancer were selected as candidate miRNAs (supplemental table 1). A total of 852 plasma samples were enrolled, including samples from 266 healthy donors and 288 patients with NSCLC in Shenzhen People’s Hospital (Shenzhen, China), as well as from 149 healthy donors and 149 patients with NSCLC in Cancer Center of Guangzhou Medical University (Guangzhou, China). The information of participators was detailed in supplemental table 2.

Screening candidate miRNA biomarkers in NSCLC

Early-screening set: Based on the S-Poly(T) Plus miRNA assay (16), qRT-PCR was first utilized to detect 486 miRNAs among three pools, which consisted of 266 healthy plasma samples, 130 patient plasma samples with NSCLC stage I and 158 samples with NSCLC stage II-IV from Shenzhen People’s Hospital. These 486 miRNAs were depicted in the volcano plot as in figure 2A and 2B. Among them, fold-change between healthy pool and stage I pool were >2 and >4 for 188 and 91 miRNAs, respectively; between stage II-IV pool were >2 and >4 for 187 and 102 miRNAs, respectively (all p values <0.05) (figure 2C). When the two results were combined, 125 significantly differentiated miRNAs (fold changes >4 and p values <0.05) (figure 2D and figure 3A and 3B) were selected and subjected to second screening.
Further-screening set: We then performed individual qRT-PCR assay to quantify each of previous identified miRNAs in five pools: Healthy (N=266), ADC Stage I (N=96), ADC Stage II-IV (N=113), SCC Stage I (N=34) and SCC Stage II-IV (N=45). miRNA targets met one of two criteria would be chosen for the next test: Stage I vs. healthy: fold-change>2, or Stage II-IV vs. healthy: fold change>2. As a result, 30 miRNAs in the ADC and 38 miRNAs in SCC were extracted and further validated in the training step (supplemental figure 1).

Training set: Candidate miRNAs obtained from the previous step were validated using each individual randomly selected from Shenzhen People’s Hospital. The number of samples was as follows: Healthy (N=20), ADC Stage I (N=20), ADC Stage II-IV (N=20), SCC Stage I (N=10) and SCC Stage II-IV (N=10). The screening criteria were identical with that in further-screening set. Consequently, 10 miRNAs in ADC and 9 miRNAs in SCC were extracted (supplemental figure 2A and 2B). However, two miRNAs (hsa-miR-183-5p and hsa-miR-144-5p) in ADC and one miRNA in SCC (hsa-miR-144-3p) were excluded because the Ct values in some individual samples were larger than 35 (marker with the square).

Validation set-1: 7 miRNAs in ADC and 9 miRNAs in SCC were validated in each individual sample from Shenzhen People’s Hospital. The expression of all miRNAs significantly differed between healthy control and patients with NSCLC Stage I or Stage II, indicating that these miRNAs could differentiate patients with NSCLC from healthy people. Therefore, 7 miRNAs were identified as potential biomarkers in ADC, and they were hsa-miR-26a-5p, hsa-miR-126-5p,
hsa-miR-139-5p, hsa-miR-152-3p, hsa-miR-451a, hsa-miR-200c-3p and
hsa-miR-3135b (figure 4A); and 9 miRNAs were identified as potential biomarkers in
SCC, including hsa-miR-26a-5p, hsa-miR-126-5p, hsa-miR-139-5p, hsa-miR-151a-3p,
hsa-miR-151a-5p, hsa-miR-151b, hsa-miR-152-3p, hsa-miR-550a-3p and
hsa-miR-3135b (figure 4B), of which, there were 5 universal biomarkers in ADC and
SCC, and they were hsa-miR-26a-5p, hsa-miR-126-5p, hsa-miR-139-5p,
hsa-miR-152-3p as well as hsa-miR-3135b.

Sensitivity and Specificity of Screening Measures

To evaluate the screening measure of potential biomarkers in NSCLC, the
sensitivity and specificity were illustrated in figure 5. In receiver operating
characteristic (ROC) curve analysis, area under ROC curve ranged from 0.64 to 0.76
in ADC ($p$ value<0.001) (figure 5A), 0.64 to 0.91 in SCC ($p$ value<0.001) (figure 5B)
and 0.67 to 0.80 for universal biomarkers ($p$ value<0.001) (figure 5C).

When the sensitivity of one single biomarker will not be sufficient for early
detection, and it has been suggested that a combination of multiple biomarkers will
potentially produce an improved ROC curve in clinical performance (20). On
multivariate logistic regression analysis (the entered variables were 7 miRNAs in
ADC, 9 miRNAs in SCC and 5 miRNAs in universal group), the statistical
significance threshold for variable entry was $p$<0.05, as results, 5/6 miRNAs as
significant predictors in ADC/SCC were remained and there were only 2 miRNAs left
in universal group. In this study, the combined miRNAs as single biomarker promised
satisfactory discrimination over diagnostic tests in ADC (5 miRNAs, AUC=0.89, $p$
value < 0.001) (figure 5A), SCC (6 miRNAs, AUC = 0.97, p value < 0.001) (figure 5B) and universal group (2 miRNAs, AUC = 0.80, p value < 0.001) (figure 5C). The maximum Youden-index (sensitivity+specificity - 1) yield a sensitivity of 98.0% and a specificity of 78.6% in ADC (cutoff = 0.0787, risk score = 0.0676 ± 0.1224 (SD) (min = -1.00 and max = 0.187), p value < 0.001); a sensitivity of 100% and a specificity of 75.7% (cutoff = 0.0075, risk score = -0.0012 ± 0.10117 (SD) (min = -1.00 and max = 0.022), p value < 0.001); and in universal group, the sensitivity and specificity were 87.5% and 94.6% (cutoff = 0.2128, risk score = 0.1987 ± 0.1129 (SD) (min = -0.90 and max = 0.23), p value < 0.001) (supplemental table 3).

Meanwhile, a formula was estimated to determine the predicted probability of being NSCLC (risk score). The developed logistic regression equation also presents the degree of contribution of each variable to the prediction. More specific, the relationship between the risk score of adenocarcinoma (P) and the relative expression of predictors is detailed as:

\[
\ln\left(\frac{P}{1-P}\right) = -6.632 \times \text{hsa-miR-26a-5p} + 36.668 \times \text{hsa-miR-126-5p} + 78.531 \times \text{hsa-miR-139-5p} - 0.042 \times \text{hsa-miR-451a} - 0.272 \times \text{hsa-miR-3135b}
\]

In squamous carcinoma, the equation is as follow:

\[
\ln\left(\frac{P}{1-P}\right) = 357.275 \times \text{hsa-miR-139-5p} + 19.103 \times \text{hsa-miR-151a-3p} - 81.391 \times \text{hsa-miR-151a-5p} + 132.895 \times \text{hsa-miR-151b-905.226} \times \text{hsa-miR-550a-3p} - 2.074 \times \text{hsa-miR-3135b-3.29}
\]

In universal group, the equation is as follow:

\[
\ln\left(\frac{P}{1-P}\right) = 87.752 \times \text{hsa-miR-139-5p} - 54.766 \times \text{hsa-miR-152-3p} - 1.651
\]
Validation of the biomarkers and risk score in NSCLC (Validation set-2)

To test the diagnostic and prognostic capability of the potential biomarkers identified in this study, 7 miRNAs in ADC and 9 miRNAs in SCC were revalidated using another set of plasma sample from Cancer Center of Guangzhou Medical University, and all the miRNAs expressed significantly different between the healthy controls and patients with Stage I and Stage II-IV (supplemental figure 3A and 3B). Therefore, 7 miRNAs in ADC and 9 miRNAs in SCC were identified as potential biomarkers for prognosis of NSCLC. In receiver operating characteristic (ROC) curve analysis, miRNAs had a much better AUC (area under ROC curve) for differentiate patients and healthy individuals than those in the validation set-1, more specific, the number ranged from 0.75 to 0.96 (p value<0.001) in ADC (supplemental figure 4A), from 0.80 to 0.99 in SCC (p value<0.001) (supplemental figure 4B) and from 0.83 to 0.98 in universal group (p value<0.001) (supplemental figure 4C).

In validation set-2, we tested prediction performance of risk score derived from validation set-1. The final prediction results revealed a sensitivity of 94.0%, a specificity of 91.6% and an overall accuracy of 92.7% in ADC; a sensitivity of 98.5%, a specificity of 51.5% and an overall accuracy of 75.0% in SCC. It was noticed that a high sensitivity produced a much lower specificity (51.5%) in SCC, more specific, 26/66 healthy control samples with continuous levels of risk score were classified into SCC group. Therefore, we considered a manually adjusted cutoff (0.14) and it might be better for prognosis with a sensitivity of 98.4%, a specificity of 90.9% and an overall accuracy of 94.5%. For universal biomarkers, the sensitivity and specificity...
were 98.7% and 17.4%, respectively; and the overall accuracy only reached 58.1%
(supplemental table 3). We therefore highlighted the separate prediction function of
risk score in ADC or SCC for greater overall accuracy.

Validation of the potential biomarker in tissue

To further determine whether our approach could identify a plasma miRNA
profile that mirrors disease progression, we analyzed the expression pattern of
potential biomarkers in non-small cell lung cancer tissues. In general, miRNAs were
highly expressed in the adjacent tissues, with lower concentrations in the cancer
tissues (figure 6). Except for hsa-miR-451a, these results were inversely correlated
with results generated from plasma profiles. Tissue samples used in this test were
collected from Shenzhen People’s Hospital (Shenzhen, China). To verify these results,
different tissue samples of NSCLC from Peking University Shenzhen Hospital
(Shenzhen, China) were tested. The higher level of miRNAs in adjacent tissues in
new samples had similar trends as in the former samples (supplemental figure 5).

Discussion

The investigation of biological and molecular features of lung cancer is critical for
identifying specific risk marker for lung cancer development, thus achieving the
earliest prediction and intervention of lung cancer (21). The recently available data
from NSCLC offered the opportunity to explore miRNA-based biomarker using
miRNA-sequence or microarray. In our study, genome-wide miRNA expression
profiling was obtained via a different platform in RT-qPCR with higher accuracy.
Moreover, our identification of aberrant circulating miRNA in the data from samples
has a number of advantages compared with previous studies. Our study also revealed
that potential biomarkers for NSCLC diagnosis could be validated in different subsets of patients. Also, we found that the area under ROC curve was higher in the validation set-2 than that in validation set-1, and the possible reason could be that the batch of samples collected from Cancer Center of Guangzhou Medical University has better quality, with perfect-matched age between patient and healthy group (supplemental table 2). Therefore, more restrictive approach and usage of different cohorts of samples could reduce the potential of misjudgment because of technical unreliability and sample heterogeneity.

Our results demonstrated these validated miRNAs can potentially serve as novel noninvasive biomarkers for early diagnosis of ADC or SCC. In the present study, we found that 11 miRNAs were dysregulated in the plasma of lung cancer patients in comparison to cancer-free control subjects, including hsa-miR-26a-5p, hsa-miR-126-5p, hsa-miR-139-5p, hsa-miR-152-3p, hsa-miR-200c-3p, has-miR-451a, hsa-miR-3135b, hsa-miR-151a-3p, hsa-miR-151a-5p, hsa-miR-151b and hsa-miR-550a-3p. As for the discrimination between ADC/SCC and healthy control, our screening strategy yielded a highest AUC of 0.89 in ADC and 0.97 in SCC, respectively. Importantly, ten miRNAs (except for hsa-miR-451a) could discriminate Stage I ADC/SCC from healthy individuals in validation set-1 as well as set-2, which generated exciting prospects for early diagnosis and prognosis. However, the accuracy of 5 universal potential biomarkers was barely satisfactory and it could be better to treat ADC and SCC separately in prognosis. According to published data, 6 out of 11 circulating miRNAs have been reported being related with NSCLC and could be potential biomarkers, which were hsa-miR-126-5p (22), hsa-miR-139-5p (23), hsa-miR-152-3p (24), hsa-miR-200c-3p (25), hsa-miR-451a (26) and hsa-miR-550a-3p (27). However, no matter in plasma/serum, tissue or cells, there are
few reports analyzing the relationship between expression levels of these four miRNAs (hsa-miR-151a-3p, hsa-miR-151a-5p, hsa-miR-151b and hsa-miR-3135b) and the occurrence of NSCLC. Multiple studies reported that hsa-miR-26a functions in cell and tissue and contributes to lung cancer development (28-30) and yet no study observed a high concordance between the circulating hsa-miR-26a and NSCLC.

It is well known that the miRNA expressions are specific for tissues, developmental stages and various diseases (31, 32). In our study, the available data herein indicated that expression patterns of miRNAs in plasma were different from those in tissue (figure 4 and figure 6). Boeri et al. also reported that circulating miRNAs had the predictive role independent from tissue specimens, because plasma miRNAs in the earlier phases of disease may be different from those required for the maintenance and the progression of the tumor (21). Teng et. al reported that miRNAs were secreted actively through small vesicles called “exosomes” that prevent them from degradation by RNases(33), and tumor exosomal miRNA might be sorted based on the miRNA biological function via a binding protein, and more specific, miRNAs with tumor suppressive effects involving cell growth suppression were encapsulated into exosomes that were subsequently released into circulation, and then miRNA levels in tumor tissue are reduced compared to that in the adjacent tissue(34).

According to this explanation, we speculated that ten miRNAs in our study, including hsa-miR-26a-5p, hsa-miR-126-5p, hsa-miR-139-5p, hsa-miR-152-3p, hsa-miR-200c-3p, hsa-miR-3135b, hsa-miR-151a-3p, hsa-miR-151a-5p, hsa-miR-151b and hsa-miR-550a-3p could be suppressive genes in non-small cell lung cancer. There have been published researches on suppressive effects of the first five miRNAs (hsa-miR-26a-5p, hsa-miR-126-5p, hsa-miR-139-5p, hsa-miR-152-3p and hsa-miR-200c-3p) in NSCLC (35-39); Ho et. al also reported that miR-550a-3p
inhibited breast cancer initiation, growth, and metastasis (40); and we need more
evidences to light the possibility of cancer suppressor for hsa-miR-3135b,
hsa-miR-151a-3p, hsa-miR-151a-5p and hsa-miR-151b. As for some exception like
hsa-miR-200c-3p, which was upregulated in four tissue samples and downregulated in
the rest of samples (figure 6), the possible reason could be that the downregualtion of
miRNA in tumor tissue was stage-dependent, and this needs further testing using
clinical samples with more details.

Presently, this study provided that the unique patterns of circulating miRNAs
might serve as noninvasive biomarkers for early diagnosis of NSCLC; however, we
were not able to investigate the pathogenesis of NSCLC when the expression of
miRNAs was changed. On the other hand, as a novel tool for diagnostic and
prognostic, the clinical utility remains to be improved. A more standardized execution
of miRNA measurements need to be established and simpler and more robust
protocols will allow potential circulating biomarkers enter the clinical application
very soon.
References


Ho JY, Hsu RJ, Wu CH, Liao GS, Gao HW, Wang TH, Yu CP. Reduced miR-550a-3p leads to breast cancer initiation, growth, and metastasis by increasing levels of ERK1 and 2. Oncotarget 2016;7:53853-68.
Figure legends

**Fig. 1** Overview of the experiment design. Flow diagram illustrates the five steps of miRNAs selection as potential biomarkers from 486 miRNAs in non-small cell lung cancer (NSCLC).

**Fig. 2** Profiling of 486 miRNAs in plasma of Non-small cell lung cancer (NSCLC). (A & B) Volcano plot illustrates miRNAs differently expressed in healthy controls compared with Stage I patients (A) and Stage II-IV patients (B); (C) fold change analysis in number of altered miRNAs; (D) 125 miRNAs dysregualted with at least 4-fold changes in control individuals compared with Stage I or Stage II-IV patients (p<0.05). miRNAs with Ct value >35 were eliminated. Data are presented as fold change ($2^{-\Delta\Delta Ct}$), and expression was normalized to a spiked-in (cel-miR-54). Samples were collected from Shenzhen People’s Hospital.

**Fig. 3** Hierarchical clustering of 125 miRNAs. Heatmap depicted miRNAs differentially expressed between healthy controls and Stage I or Stage II-IV patients in ADC (A) and SCC (B). Samples were collected from Shenzhen People’s Hospital.

**Fig. 4** Validation of candidate miRNAs in each individual of samples from Shenzhen People’s Hospital. (A) 209 healthy donors and 209 patients with ADC; (B) 79 healthy donors and 79 patients with SCC. Data are shown as means ± SE, ns= not significant, *p<0.05 and ***p<0.001. Samples were collected from Shenzhen People’s Hospital.

**Fig. 5** Receiver operating characteristic curves (ROC). (A) ROC of 7 individual miRNAs and 5 miRNAs as one combined biomarker in ADC; (B) ROC of 9 individual miRNAs and 6 miRNAs as one combined biomarker in SCC; (C) ROC of 5 individual
miRNAs as potential universal biomarker in all samples, as well as 2 miRNAs as one combined biomarker. Samples were obtained from Shenzhen People’s Hospital.

Fig. 6 Expression levels of 11 miRNAs in adjacent tissues and cancer tissues. Samples were collected from Shenzhen People’s Hospital. It was paired two-tailed Student’s t test that used in comparing the miRNA expressions between the tumor tissues and adjacent tissues. ns=not significant, *p<0.05, **p<0.01 and ***p<0.001.
**Early-screening set:** S-poly(T)Plus qPCR for 486 miRNAs on two pooled samples  
*Screening criteria:* Stage I vs. healthy: fold-change $>4$ or Stage II-IV vs. healthy: fold change $>4$ and $P<0.05$  
130 NSCLC I 158 NSCLC II-IV 266 normal controls (NC)

**Further-screening set:** S-poly(T)Plus qPCR for 125 miRNAs on five pooled samples  
*Screening criteria:* Stage I vs. healthy: fold-change $>2$, or Stage II-IV vs. healthy: fold change $>2$ and $P<0.05$  
96 ADC I 113 ADC II-IV 34 SSC I 45 SSC II-IV 266 NC

**Training set:** S-poly(T)Plus qPCR of 30+38 miRNAs in small number of individual samples  
*Screening criteria:* Stage I vs. healthy: fold-change $>2$, or Stage II-IV vs. healthy: fold change $>2$ and $P<0.05$  
20 ADC I 20 ADC II-IV 10 SSC I 10 SSC II-IV 30 NC

**Validation set-1:** S-poly(T)Plus qPCR of 7+9 miRNAs in individual samples  
288 NSCLC (Shenzhen) 266 NC

**Validation set-2:** S-poly(T)Plus qPCR of 7+9 miRNAs in individual samples  
149 NSCLC (Guangzhou) 149 NC

*Candidate miRNAs*  

---

*Figure 1*
Figure 3
Figure 4
Cancer Epidemiology, Biomarkers & Prevention

Circulating plasma microRNAs as potential biomarkers of non-small cell lung cancer obtained by high-throughput real-time PCR profiling

Yanqin Niu, Mingyang SU, Yike Wu, et al.

Cancer Epidemiol Biomarkers Prev  Published OnlineFirst October 30, 2018.

Updated version  Access the most recent version of this article at: doi:10.1158/1055-9965.EPI-18-0723

Supplementary Material  Access the most recent supplemental material at: http://cebp.aacrjournals.org/content/suppl/2018/10/30/1055-9965.EPI-18-0723.DC1

Author Manuscript  Author manuscripts have been peer reviewed and accepted for publication but have not yet been edited.

E-mail alerts  Sign up to receive free email-alerts related to this article or journal.

Reprints and Subscriptions  To order reprints of this article or to subscribe to the journal, contact the AACR Publications Department at pubs@aacr.org.

Permissions  To request permission to re-use all or part of this article, use this link http://cebp.aacrjournals.org/content/early/2018/10/30/1055-9965.EPI-18-0723. Click on "Request Permissions" which will take you to the Copyright Clearance Center's (CCC) Rightslink site.