

Early detection of squamous cell lung cancer in sputum by a panel of microRNA markers

Lingxiao Xing¹, Nevins W Todd², Lei Yu¹, HongBin Fang³ and Feng Jiang¹

¹Department of Pathology, University of Maryland School of Medicine, Baltimore, MD, USA; ²Department of Medicine, University of Maryland School of Medicine, Baltimore, MD, USA and ³Division of Biostatistics of The University of Maryland Greenebaum Cancer Center, University of Maryland School of Medicine, Baltimore, MD, USA

Squamous cell carcinoma is a common form of lung cancer, the leading cause of cancer deaths in the world. Identifying early stage lung squamous cell carcinoma patients who would benefit most from effective therapies will reduce the mortality. We have previously shown that microRNAs (miRNAs) were stably present in sputum and potentially useful in diagnosis of lung cancer. The objective of this study was to develop a panel of miRNAs that can be used as a sputum-based test for early stage squamous cell carcinoma of the lungs. This study contained three phases: (1) marker discovery by profiling miRNA expression signatures on 15 lung squamous cell carcinoma and matched normal lung tissue samples with GeneChip miRNA Array; (2) marker optimization by real-time quantitative RT-PCR on sputum of a case–control cohort of 48 stage I lung squamous cell carcinoma patients and 48 healthy individuals; and (3) marker validation on an independent set including 67 lung squamous cell carcinoma patients and 55 healthy subjects. On the surgical tissues, six miRNAs were identified, of which three were overexpressed and three were underexpressed in all 15 tumors. On the sputum samples of the case–control cohort, three (*miR-205*, *miR-210* and *miR-708*) of the six miRNAs were selected, which in combination produced the best prediction in distinguishing lung squamous cell carcinoma patients from normal subjects with 73% sensitivity and 96% specificity. Validation of the marker panel in the independent populations confirmed the sensitivity and specificity that provided a significant improvement over any single one alone. The sputum markers showed the potential to improve the early detection of lung squamous cell carcinomas.

Modern Pathology (2010) **23**, 1157–1164; doi:10.1038/modpathol.2010.111; published online 4 June 2010

Keywords: biomarkers; early detection; lung cancer; microRNA; sputum

Non-small-cell lung cancer is the number one cancer killer in the United States and worldwide.¹ Non-small-cell lung cancer comprises three major histological subtypes: squamous cell carcinoma, adenocarcinoma and large cell carcinoma. Squamous cell carcinoma is usually found in the central airways, whereas AC and squamous cell carcinoma are often peripherally located in the lungs. The overall 5-year survival rate for stage I non-small-cell lung cancer patients who are typically treated with surgery remain up to 80%. In contrast, only 5–15% and <2% of patients with stage III and IV non-small-cell lung cancer are alive after 5 years.¹ The

ability to identify early stage lung cancer patients who would benefit most from effective therapies will reduce the mortality. Chest X-ray has been used for its early detection, however, the sensitivity is low.² Computed tomography (CT) provides excellent anatomic information and has an increasing role in noninvasive diagnosis of early non-small-cell lung cancer.^{2,3} However, it has limited ability to differentiate between benign and malignant lesions for tumors, particularly centrally located cancer.⁴ Therefore, the development of new noninvasive approaches that can be used alone or complement CT imaging in more accurately identifying squamous cell carcinoma, which is predominantly located in central airways, is clinically important.

As morphological changes of exfoliated bronchial epithelium from sputum are associated with incident squamous cell carcinoma of the lungs,^{5,6} cytologic analysis of sputum has been used clinically for its diagnosis.⁶ However, it was no more

Correspondence: Dr F Jiang, Department of Pathology, University of Maryland School of Medicine, 10 South Pine Street, MSTF 7th floor, Baltimore, MD 21201-1192, USA.

E-mail: fjiang@som.umaryland.edu

Received 17 January 2010; revised 16 April 2010; accepted 19 April 2010; published online 4 June 2010

effective than chest radiographs in detecting lung cancer in several large prospective randomized trials.³ The molecular genetic changes detected in sputum may correlate with those exist in lung tumors, and could occur before morphological changes that can be found by a cytological test.^{6–8} Therefore, molecular study of sputum has been suggested to be more sensitive than cytology in diagnosis of lung cancer.^{6–10} For instance, hypermethylation of *p16* gene was found in sputum collected from patients with lung squamous cell carcinomas, 5–35 months before sputum cytological and clinical diagnoses.¹¹

MicroRNAs (miRNAs) are a newly discovered class of small noncoding RNAs that have critical roles in a wide spectrum of biologic and pathologic processes.¹² Furthermore, miRNAs are emerging as tissue-specific biomarkers with potential for clinical applicability in indentifying and defining cancer type.^{13–15} Our recent proof of principle study¹⁶ showed that endogenous miRNAs were present in sputum in a remarkably stable form and could reliably be measured by reverse transcription (RT)-quantitative PCR (qPCR). In addition, detecting elevated expression of a single miRNA, *miR-21*, produced a higher sensitivity in diagnosis of lung cancer compared with sputum cytology. Our data suggested that the measurement of altered miRNA expressions in sputum could be a useful noninvasive approach for lung cancer diagnosis. However, the sensitivity reached by a single miRNA is too low for clinical application.¹⁶

It has been proven that lung squamous cell carcinoma is a heterogeneous disease and develops from complex and multistep processes.^{17,18} We therefore hypothesized that simultaneous assessment of a panel of tumor-specific miRNAs that, used in combination in sputum, could provide a sensitive and specific diagnostic test for early squamous cell carcinoma. To verify the hypothesis, we first identified miRNA signatures of stage I lung squamous cell carcinoma using miRNA profiling on primary tumor tissues. From these signatures, we then optimized and validated a panel of miRNAs that could be detected in sputum for early detection of lung squamous cell carcinoma with an acceptable diagnostic sensitivity and specificity.

Materials and methods

Patients and Clinical Specimens

To define miRNA signatures for lung squamous cell carcinoma, surgical specimens were obtained from 15 lung cancer patients who had either a lobectomy or a pneumonectomy between 1 March 2000 and 28 June 2003 at the University of Maryland Medical Center. All cases were diagnosed with histologically confirmed stage I squamous cell carcinoma of the lungs. None of the patients had received preopera-

Table 1 Characteristics of lung squamous cell carcinoma patients and healthy subjects in an independent cohort

	Lung squamous cell carcinoma (n = 67)	Controls (n = 55)	P-value
Age	68 (s.d., 12.5)	65 (s.d., 10.5)	>0.05
Sex			
Female	21	20	>0.05
Male	46	35	>0.05
Race			
White	54	43	>0.05
African American	13	12	>0.05
Smoking status			
Pack-years	31.1 (s.d., 23.5)	26.8 (s.d., 27.6)	>0.05
Stage			>0.05
I	14		
II	16		
III	19		
IV	18		

tive adjuvant chemotherapy or radiotherapy. Tumor tissues were intraoperatively dissected from the surrounding lung parenchyma; paired normal lung tissues were also obtained from the same patients at an area distant from their tumors. Serial cryostat sections from the specimens were stained with hematoxylin and eosin to confirm the diagnosis based on the most recent WHO classification of tumors of the lung.¹⁹

To optimize a panel of miRNAs that could be detected in sputum, 48 stage I lung squamous cell carcinoma patients and an equal number of normal subjects were recruited. The case and control were matched in the ratio of 1:1 by age, gender and smoking history as a case–control cohort (Supplementary Table 1). Sputum was collected from the participants as described in our recent reports.^{16,20–22} To further validate the identified sputum markers, we collected sputum specimens from an independent set of 67 lung squamous cell carcinoma patients and 55 healthy controls. The demographic and clinical characteristics of the cancer patients are summarized in Table 1. The study was approved by Institutional Review Board.

RNA Isolation

Total RNA containing small RNA was extracted from tissue and sputum specimens as described in our previous study¹⁶ by using a mirVana miRNA Isolation Kit (Ambion, Austin, TX, USA). The purity and concentration of RNA were determined from OD260/280 readings using a dual-beam UV spectrophotometer (Eppendorf AG, Hamburg, Germany). RNA integrity was determined by capillary electrophoresis using the RNA 6000 Nano Lab-on-a-Chip kit and the Bioanalyzer 2100 (Agilent Technologies, Santa

Clara, CA, USA). Only RNA extracts with RNA integrity number values >6 underwent in further analysis.

miRNA Profiling of Surgical Resected Lung Squamous Cell Carcinoma Tissues

miRNA profiling was performed by using GeneChipR miRNA Array (Affymetrix, Santa Clara, CA, USA). The array comprised 7815 probe sets. Of the probes, 6703 encoded miRNA coverage of human, mouse, rat, canine and rhesus macaque, 922 encompassed human snoRNAs and scaRNAs. In all, 190 control targets were also included in the array that contained 95 background probe sets, 63 hybridization control probe sets, 22 oligonucleotide spike-in control probe sets and 10 identical probes. Microarray experiments were conducted with all matched malignant and normal sample pairs according to the manufacturer's instructions. Briefly, 3 μ g total RNA was labeled with Biotin FlashTag Biotin Labeling Kit (Affymetrix). The labeling reaction was hybridized on the miRNA Array in Affymetrix Hybridization Oven 640 (Affymetrix) at 48 °C for 16 h. The arrays were stained with Fluidics Station 450 using fluidics script FS450_0003 (Affymetrix), and then scanned on an Axon 4000B microarray scanner (Axon Instruments, Foster City, CA, USA). miRNA probe outliers were defined as per the manufacturer's instructions (Affymetrix) and further analyzed for data summarization, normalization and quality control by using the web-based miRNA QC Tool software (www.affymetrix.com). To find miRNAs that were statistically differentially expressed between squamous cell carcinoma specimens and corresponding normal tissues, we expected the fold difference between lung tumor and normal samples to be >1.0 with <0.01 statistical significance. We analyzed the normalized microarray data by using GenePattern (<http://www.broad.mit.edu>), BRB-ArrayTools version 3.6 (<http://linus.nci.nih.gov/BRB-ArrayTools.html>), and microarray software suite 4 (TM4) (<http://www.tm4.org>). Finally, we performed tree visualization by using Java Treeview 1.0 (Stanford University School of Medicine, Stanford, CA, USA).

miRNA Quantitative Real-Time RT-PCR

The identified miRNAs were evaluated in sputum by using RT-qPCR with Taqman miRNA assays (Applied Biosystems, Foster City, CA, USA) as previously described.¹⁶ Ct values of the target miRNAs were normalized in relation to that of small nuclear *U6 RNA*. *U6 RNA* was proven as an internal control for miRNA quantification in sputum in our previous study.¹⁶ All assays were performed in triplicates, and one no-template control and two interplate controls were carried along in each experiment. Expression levels of the miRNAs were

calculated using comparative Ct method as previously described.¹⁶

To determine the sensitivity and dynamic range of miRNA quantification in sputum, RNA was extracted from ten sputum specimens and then diluted at different orders of magnitude in diethylpyrocabonate (DEPC) water (Sigma Chemical, St Louis, MO, USA). Expressions of the miRNAs were then assessed by using RT-qPCR in the samples described above. All tests were performed in triplicates.

Statistical Analysis

Statistical analysis of RT-qPCR data were conducted using Statistical Analysis System software version 6.12 (SAS Institute, Cary, NC, USA). Spearman rank correlation was carried out to analyze the correlation between the expressions of the identified miRNAs. All *P*-values shown were two-sided, and a *P*-value of <0.05 was considered statistically significant. Receiver-operator characteristic (ROC) curve analysis was undertaken using expression level for each miRNA in sputum from cancer patients and healthy controls by Analyse-it software (Analyse-it Software, Leeds, UK). Logistic regression was used to generate prediction model building. For each validated miRNA biomarker, we constructed the ROC curve and computed the area under ROC (AUC) value by numerical integration of the ROC curve. Validated biomarkers were fitted into logistic regression models, and the stepwise backward model selection was performed to determine the best discriminating combinations of miRNAs. Furthermore, contingency table and logistic regression analysis were applied to determine the associations between expression levels of the miRNAs and both clinicopathologic and demographic characteristics of the cases and controls.

Results

Identifying miRNA Signatures Whose Aberrant Expression Levels were Specific to Squamous Cell Carcinoma of the Lung

We used GeneChipR miRNA Arrays to profile expression signatures of mature miRNAs on lung squamous cell carcinoma tissues. As the miRNA Array comprised probe sets for miRNA coverage of human, mouse, rat, canine and rhesus macaque, and human snoRNAs and scaRNAs as well, we only analyzed and compared expressions of 818 human mature miRNAs in the tumor and normal tissue specimens. When *P*-value <0.01 was used as a cutoff, of the human miRNAs, 21 were overexpressed and 26 were underexpressed with ≥ 1.0 fold-change in cancer group ($P < 0.01$) (Figure 1 and Supplementary Table 2). Using a predefined criterion of a change \geq twofold, we identified six miRNAs

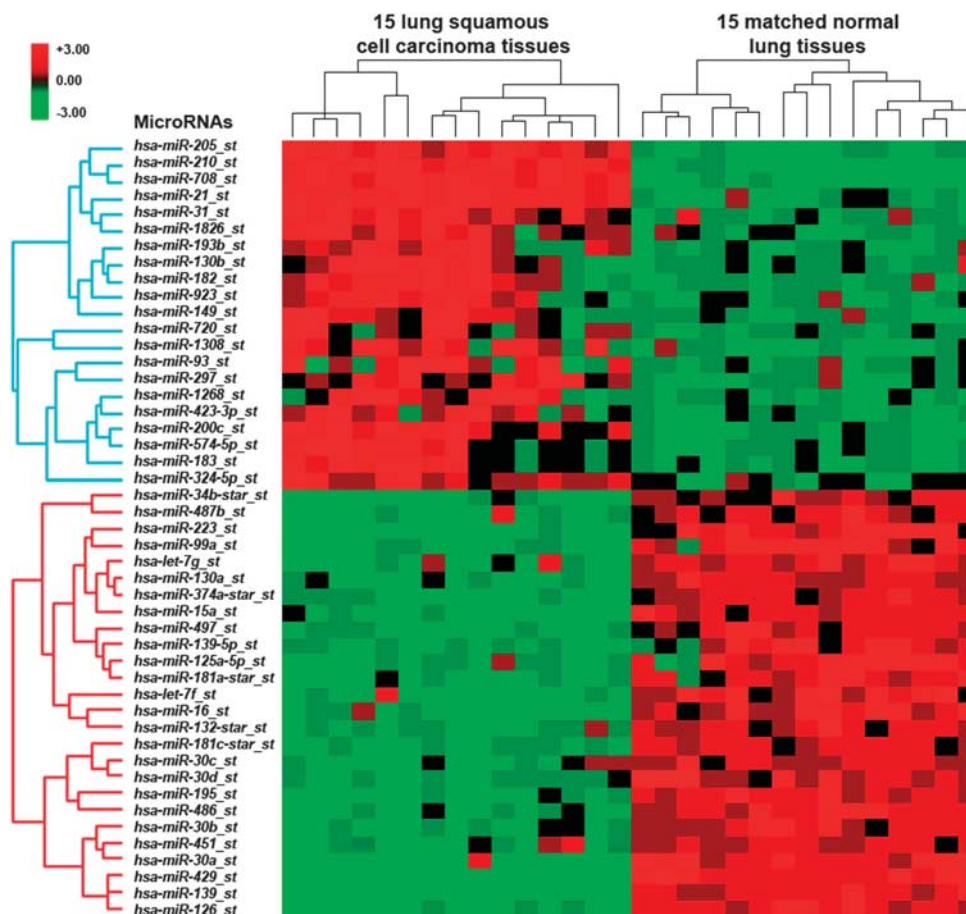


Figure 1 miRNAs differentially expressed in lung squamous cell carcinoma versus normal lung tissues. Hierarchical clustering of 47 miRNA genes with a significantly different expression ($P < 0.01$) in tumor tissues. Rows represent individual genes; columns represent individual tissue samples. The scale represents the intensity of gene expression (\log_2 scale ranges between -3.0 and 3.0).

that were statistically differently expressed between the paired tumor and normal samples (all $P < 0.01$). These included three miRNAs (*miR-205*, *miR-210* and *miR-708*) that were overexpressed, and three miRNAs (*miR-126*, *miR-139* and *miR-429*) that were underexpressed in tumor specimens. It should be noted that altered expressions of the six miRNAs existed in all 15 lung squamous cell carcinoma tissues compared with the paired normal specimens. Therefore, the six miRNAs proceed to the next phase of the study.

Optimizing a Panel of Highly Specific and Sensitive Sputum miRNA Markers for Squamous Cell Carcinoma of the Lung

To determine if the six newly identified miRNAs from surgical tissues by the GeneChipR miRNA Array could be reliably detected in sputum by RT-qPCR, we first prepared two RNA pools containing equal amounts of RNA from sputum samples of 10 cancer patients and 10 cancer-free individuals, respectively. Expression of each miRNA was then measured on the pooled RNA samples. All tested

miRNAs had ≤ 32 Ct values in both pools, indicating that expression of the miRNAs could easily be determined in sputum. To further decide the sensitivity of detecting the miRNAs by RT-qPCR in sputum, the total RNA was diluted in DEPC water at serial concentrations. The diluted RNAs served as experimental samples for measuring expression of each miRNA. The assay had a dynamic range of at least eight orders of magnitude ($R^2 = 0.992$), and was capable of detecting as little as 10 copies of the target genes. In addition, there was an excellent linearity between the RNA input and the Ct values for the miRNA tested (data not shown). Altogether, the miRNAs identified from the primary tumor tissues can readily be detectable in sputum. The six miRNAs were therefore continually tested in all individual sputum samples collected from 48 patients diagnosed with stage I lung squamous cell carcinoma and 48 healthy subjects.

As shown in Table 2, *miR-205*, *miR-210* and *miR-708* showed higher expression levels, whereas *miR-126*, *miR-139* and *miR-429* showed lower expression levels in cancer patients' sputum compared with sputum of cancer-free individuals (all $P < 0.05$). The data produced from the sputum analysis was in

Table 2 Expression levels of the six miRNAs and their diagnostic significance in sputum of 48 lung squamous cell carcinoma patients and 48 healthy controls

miRNA	Mean (s.d.) in cancer patients	Mean (s.d.) in healthy controls	P ^a	AUC ^a (s.e.)	Cutoffs	Sensitivity (%)	Specificity (%)
miR-205	4.15 (1.87)	15.06 (4.92)	<0.001	0.789 (0.039)	9.12	65	90
miR-210	0.56 (0.23)	2.64 (0.84)	<0.001	0.748 (0.022)	1.64	58	79
miR-708	1.86 (1.05)	6.01 (1.98)	<0.001	0.756 (0.028)	4.48	65	75
miR-126	6.37 (2.12)	0.86 (0.06)	<0.001	0.702 (0.027)	3.23	60	79
miR-139	0.03 (0.01)	0.01 (0.01)	0.026	0.659 (0.023)	0.02	56	77
miR-429	0.22 (0.09)	0.08 (0.01)	0.021	0.623 (0.016)	0.13	60	73

Abbreviation: AUC, the area under receiver operating characteristic curve.

^aBoth P-value and AUC were obtained using the U6-normalized values.

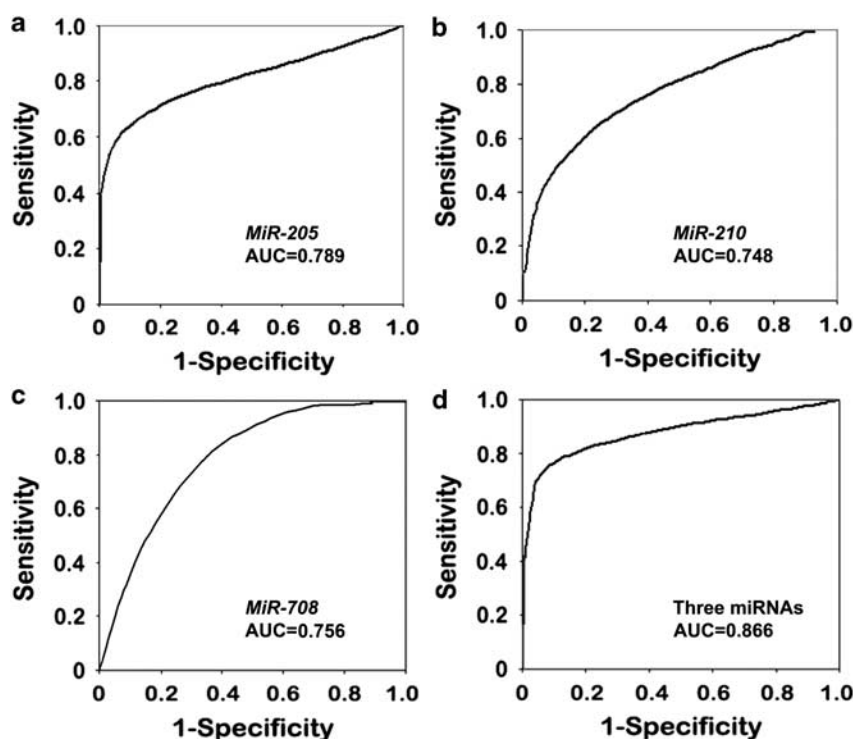


Figure 2 ROC curve analysis of expression levels of the three miRNAs in sputum of 48 patients diagnosed with stage I lung squamous cell carcinoma and 48 healthy subjects. The area under the ROC curve (AUC) for each miRNA conveys its accuracy for differentiation of lung squamous cell carcinoma patients and healthy subjects in terms of sensitivity and specificity. The individual genes produced 0.748–0.789 AUC values (a–c), being significantly lower than 0.866 AUC by the three genes combined as a marker panel (d) (all $P < 0.05$).

agreement with the results obtained from the tissue specimens. ROC analyses were then performed to evaluate the capability of using the miRNAs to discriminate between cancer patients and healthy individuals. As depicted in Table 2, the six miRNAs showed 0.623–0.789 AUC values. When optimum cutoffs were selected, the miRNAs yielded 56–65% sensitivity and 73–90% specificity. Of the genes, miR-205 was the best single miRNA with 0.79 AUC, resulting in 65% sensitivity and 90% specificity. The findings imply that the identified miRNAs held promise as sputum markers for squamous cell carcinoma of the lungs.

To optimize a small panel of miRNA markers for diagnosis of lung squamous cell carcinoma with

high sensitivity and specificity, logistic regression of all six miRNAs using a backward elimination approach was performed. One of the logistic regression models was built based on three miRNAs, miR-205, miR-210 and miR-708, which in combination provided the best prediction. The combination of the three miRNAs produced 0.866 AUC, being considerably higher than 0.623–0.789 AUC values of each individual gene in distinguishing cancer patients from normal subjects (all $P < 0.05$) (Figure 2). Accordingly, given a specificity of 96%, the composite panel of the miRNAs revealed a sensitivity of 73%, which was significantly higher than 65% sensitivity of miR-205, the single best miRNA in detection of lung squamous cell

carcinoma ($P < 0.05$) (Supplementary Figure 1). Furthermore, Spearman rank correlation analysis showed that the estimated correlations among expression levels of the miRNAs in sputum were low (all $R^2 < 0.50$, $P > 0.05$). The data suggested that expressions of the three miRNAs were complementary to each other, further supporting that the combined analyses of the genes outperformed a single one used alone. Therefore, the optimal panel of the miRNAs provides high diagnostic efficiency for early lung squamous cell carcinoma in sputum.

Validating the Sputum miRNA Markers in an Independent Set of Squamous Cell Carcinoma Patients

To further evaluate the diagnostic performance of the optimal panel of markers, the three miRNAs were assessed on sputum samples collected from 67 patients with different stages of squamous cell carcinoma and 55 healthy controls. The miRNAs showed higher expression levels in the squamous cell carcinoma patients compared with the healthy controls (all $P < 0.001$) (Supplementary Table 3). However, the mean expression level of each miRNA among the healthy subjects was not significantly different from that of the healthy individuals in the above case-control cohort. Similarly, the mean levels of the genes among the squamous cell carcinoma patients in the validation set did not differ from those of the cancer patients in the case-control cohort (all $P > 0.05$, respectively). Using cutoffs that maximized the sum of sensitivity and specificity on the case-control data, the panel of the markers produced 72% sensitivity and 95% specificity for lung squamous cell carcinoma in the independent validation cases. The parameters were similar to those (73 and 96%) in the case-control cohort that only consisted of stage I squamous cell carcinoma patients (all $P > 0.05$). Furthermore, the markers had similar sensitivity in the diagnosis of stage I, II, III and IV squamous cell carcinomas ($P > 0.05$), while maintaining 95% specificity (Table 3). The finding that altered expressions of the miRNAs are found not only in advanced stage, but also in early stage of squamous cell carcinoma is

Table 3 Diagnostic performance of the miRNA marker panel in a cohort of 67 lung squamous cell carcinoma patients with different stages and 55 healthy subjects

	Sensitivity, %	Specificity, %
Cases with all stages	72 (48/67)*	95 (52/55)
<i>Cases with different stages</i>		
I	71 (10/14)*	95 (52/55)
II	69 (11/16)*	95 (52/55)
III	74 (14/19)*	95 (52/55)
IV	72 (13/18)*	95 (52/55)

* $P > 0.05$.

an important characteristic if they are to be used for early detection. Moreover, there was no association of expression levels of the three miRNAs with the age, gender, ethnic group, tumor stage or histories of smoking of the lung cancer patients and normal individuals (Supplementary Tables 4 and 5) (all $P > 0.05$). Taken together, the results confirm that the optimal set of miRNAs could be used as promising biomarkers for the early detection of lung squamous cell carcinoma.

Discussion

In this study, we applied a systematic approach by using different techniques in three individual populations to discover, optimize and validate miRNA biomarkers for early stage lung squamous cell carcinoma. The developed sputum-based marker panel yielded a sensitivity of 73% and a specificity of 96%. This study extends our previous research efforts to develop noninvasive or minimally invasive diagnostic means for lung cancer.^{16,20-22} Given the expenses associated with quantitative molecular analysis of multiple genes by a RT-qPCR platform, a marker panel with the smallest number of miRNAs and highest diagnostic accuracy would provide a cost-effective assay for squamous cell carcinoma of lung cancers.

Among the three miRNAs identified, *miR-205* showed the best prediction for lung squamous cell carcinoma. Although its biologic role of *miR-205* in squamous cell carcinoma remains to be analyzed, upregulation of *miR-205* may participate in determining a squamous phenotype and the development and progression of squamous cell carcinoma.¹⁵ For instance, *miR-205* was identified as one of six miRNAs that highly expressed in squamous cell carcinoma compared with adenocarcinoma of the lungs.¹⁴ Furthermore, overexpression of *miR-205* was found in head and neck squamous cancer cell lines.^{23,24} In addition, *miR-205* was suggested as a marker of stratified squamous epithelium.^{25,26} Most recently, Lebanony *et al*¹⁵ showed that measuring *miR-205* expression on surgical tumor tissues produced a sensitivity of 96% and a specificity of 90% in the identification of squamous cell carcinoma of the lungs. In good agreement with the findings, our present data showed that *miR-205* was overexpressed in lung squamous cell carcinoma, and more importantly, assessment of its expression in sputum showed reasonable accuracy in diagnosis of early stage squamous cell carcinoma. Therefore, *miR-205* could be a highly specific marker for lung cancer of squamous histology. *miR-210* can regulate the hypoxic response of tumor cells and tumor growth.²⁷ Elevated expression of *miR-210* was significantly associated with aggressiveness of lymph node-negative, estrogen receptor-positive human breast cancer.²⁸ Moreover, detection of increased *miR-210* expression level in serum

could be useful as one of miRNA markers for patients with diffuse large B-cell lymphoma and pancreatic ductal adenocarcinoma.^{29,30} Our present observation is consistent with the previous findings, and further suggests that *miR-210* might be another important sputum-based marker for lung squamous cell carcinoma. *miR-708* was recently suggested as one of nine miRNAs whose overexpressions in tumor tissues were associated with recurrence of stage I non-small-cell lung cancer after surgical resection. Upregulation of *miR-708* might lead to lung cancer development and progression.³¹ Our data support the previous finding³¹ that the elevated expression level of *miR-708* is common in lung tumor tissues, and further show that detection of *miR-708* expression in sputum could diagnose early squamous cell carcinoma of the lungs. Our primary goal of this study is marker development. We show for the first time that measuring altered expressions of a panel of miRNAs in sputum might be a potential noninvasive test for early lung squamous cell carcinoma. Further study is ongoing in our laboratory to analyze the biological relevance of dysfunction of the miRNAs in lung tumorigenesis and determine whether the markers could also identify lung adenocarcinoma with high sensitivity and specificity.

A majority of the previously identified lung cancer associated molecular genetic changes were related to the smoking status. Some of the changes can be found in healthy smokers who never develop lung cancer.³² The use of such molecular genetic alterations as biomarkers might produce false-positive diagnostic rate or over-diagnosis, thus impeding their application in clinical settings in screening or early detection of lung cancer. The miRNAs identified from the present research is fairly encouraging as diagnostic markers, because they function independently of subject age, gender or ethnic subgroup. In particular, their expression levels do not relate with number of smoking packer-years of both cancer patients and healthy subjects. The observation is supported by recent reports,^{16,33} in which altered expressions of some miRNAs might not be associated with smoking molecular damage in bronchial epithelium. Therefore, the identified miRNAs could dysregulate in a cancer-specific manner, although the findings need to be further confirmed. Moreover, no significant differences of the miRNA expression levels were observed for the cancerous samples at different stages of the disease, implying that the potential markers were not stage specific. The results further provide evidence that this marker panel could be useful for the early detection of lung squamous cell carcinoma.

Although the identified panel of markers could distinguish lung squamous cell carcinoma patients from healthy individuals with high accuracy, the major limitation of this study is that it did not include sputum samples from patients with other benign or reactive lung lesions, in particular,

patients with chronic obstructive pulmonary disease. Chronic obstructive pulmonary disease may have some effects on the miRNA profile in sputum.^{34–36} In an ongoing study, we are currently collecting sputum from chronic obstructive pulmonary disease patients, who are matched with lung squamous cell carcinoma patients for smoking, gender, age and pulmonary function. We will then test the miRNAs on the sputum specimens and compared the results with those obtained from the cancer patients to further evaluate diagnostic performance of the marker pane in differentiating squamous cell carcinomas from other benign lesions, especially, chronic obstructive pulmonary disease.

CT provides excellent anatomic imaging, but has limitations in uncertain rate for central tumors that are mainly squamous cell carcinoma histological type.⁴ This justifies the requirement for a new diagnostic technique of such lesions with suspicious imaging features. If molecular markers with high specificity and reasonable sensitivity for cancers that will develop can be identified, such markers could be combined with CT to screen high-risk individuals, allowing molecular detection and visualization of clinically relevant early lesions. This would greatly minimize unnecessary and potentially life-threatening procedures in patients with benign lesions. In this study, we showed that the composite panel of the miRNAs revealed a sensitivity of 73% and a specificity of 96%. The specificity in identifying central lung squamous cell carcinoma is fairly high. The findings have important clinical significance, because given the high specificity, integration of the sputum-based assay with CT in future would overcome the weakness of the imaging analysis, by which over-diagnosis of central lung squamous cell carcinoma is common. Furthermore, taking advantage of outstanding anatomic resolution by CT, the integrated approach could surmount major technical obstacle of sputum-based molecular diagnosis, by which, it is difficult to localize the specific source of the abnormal cells. Moreover, future evaluation of performance of the marker panel for identification of stage 0 lung squamous cell carcinoma patients or screening for the disease in high-risk populations is required.

In conclusion, we have developed a panel of miRNAs that can be reliably measured in sputum. The assessment of the miRNAs could be used as a noninvasive and cost-effective diagnostic tool for early squamous cell lung cancer. Nonetheless, a large multi-center clinical project to further validate the full utility in large prospective cohorts is warranted before it could potentially be adopted in routine clinical settings.

Disclosure/conflict of interest

The authors declare no conflict of interest.

References

- 1 Higgins MJ, Ettinger DS. Chemotherapy for lung cancer: the state of the art in 2009. *Expert Rev Anticancer Ther* 2009;9:1365–1378.
- 2 Frost JK, Ball Jr WC, Levin ML, *et al*. Early lung cancer detection: results of the initial (prevalence) radiologic and cytologic screening in the Johns Hopkins study. *Am Rev Respir Dis* 1984;130:549–554.
- 3 Hirsch FR, Franklin WA, Gazdar AF, *et al*. Early detection of lung cancer: clinical perspectives of recent advances in biology and radiology. *Clin Cancer Res* 2001;7:5–22.
- 4 Toloza EM, Harpole L, McCrory DC. Noninvasive staging of non-small cell lung cancer: a review of the current evidence. *Chest* 2003;123:137–146.
- 5 Saccomanno G, Archer VE, Auerbach O, *et al*. Development of carcinoma of the lung as reflected in exfoliated cells. *Cancer* 1974;33:256–270.
- 6 Thunnissen FB. Sputum examination for early detection of lung cancer. *J Clin Pathol* 2003;11:805–810.
- 7 Belinsky SA. Gene-promoter hypermethylation as a biomarker in lung cancer. *Nat Rev Cancer* 2004;4:707–717.
- 8 Li R, Todd NW, Qiu Q, *et al*. Genetic deletions in sputum as diagnostic markers for early detection of stage I non-small cell lung cancer. *Clin Cancer Res* 2007;13:482–487.
- 9 Yakubovskaya MS, Spiegelman V, Luo FC, *et al*. High frequency of K-ras mutations in normal appearing lung tissues and sputum of patients with lung cancer. *Int J Cancer* 1995;63:810–814.
- 10 Mao L, Hruban RH, Boyle JO, *et al*. Detection of oncogene mutations in sputum precedes diagnosis of lung cancer. *Cancer Res* 1994;1:1634–1637.
- 11 Belinsky SA, Liechty KC, Gentry FD, *et al*. Promoter hypermethylation of multiple genes in sputum precedes lung cancer incidence in a high-risk cohort. *Cancer Res* 2006;66:3338–3344.
- 12 Ambros V. MicroRNA pathways in flies and worms: growth, death, fat, stress, and timing. *Cell* 2003;6:673–676.
- 13 Lu J, Getz G, Miska EA, *et al*. MicroRNA expression profiles classify human cancers. *Nature* 2005;435:834–838.
- 14 Yanaihara N, Caplen N, Bowman E, *et al*. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* 2006;9:189–198.
- 15 Lebanony D, Benjamin H, Gilad S, *et al*. Diagnostic assay based on hsa-miR-205 expression distinguishes squamous from nonsquamous non-small-cell lung carcinoma. *J Clin Oncol* 2009;27:2030–2037.
- 16 Xie Y, Todd NW, Liu Z, *et al*. Altered miRNA expression in sputum for diagnosis of non-small cell lung cancer. *Lung Cancer* 2010;67:170–176.
- 17 Colby TV, Wistuba II, Gazdar A. Precursors to pulmonary neoplasia. *Adv Anat Pathol* 1998;5:205–209.
- 18 Minna JD, Roth JA, Gazdar AF. Focus on lung cancer. *Cancer Cell* 2002;1:49–52.
- 19 Hammar SP, Brambilla C, Pugatch B, *et al*. Squamous cell carcinoma. In: Travis WD, Brambilla E, Muller-Hermelink HK, *et al* (eds). *Pathology and Genetics, Tumours of the Lung, Pleura, Thymus and Heart*. IARC Press: Lyon, France, 2005, pp 26–30.
- 20 Katz RL, Zaidi TM, Fernandez RL, *et al*. Automated detection of genetic abnormalities combined with cytology in sputum is a sensitive predictor of lung cancer. *Mod Pathol* 2008;8:950–960.
- 21 Qiu Q, Todd NW, Li R, *et al*. Magnetic enrichment of bronchial epithelial cells from sputum for lung cancer diagnosis. *Cancer* 2008;114:275–283.
- 22 Jiang F, Todd NW, Qiu Q, *et al*. Combined genetic analysis of sputum and computed tomography for noninvasive diagnosis of non-small-cell lung cancer. *Lung Cancer* 2009;66:58–63.
- 23 Jiang J, Lee EJ, Gusev Y, *et al*. Real-time expression profiling of microRNA precursors in human cancer cell lines. *Nucleic Acids Res* 2005;33:5394–5403.
- 24 Tran N, McLean T, Zhang X, *et al*. MicroRNA expression profiles in head and neck cancer cell lines. *Biochem Biophys Res Commun* 2007;358:12–17.
- 25 Ryan DG, Oliveira-Fernandes M, Lavker RM. MicroRNAs of the mammalian eye display distinct and overlapping tissue specificity. *Mol Vis* 2006;12:1175–1184.
- 26 Gregory PA, Bert AG, Paterson EL, *et al*. The miR-200 family and miR-205 regulate epithelial to mesenchymal transition by targeting ZEB1 and SIP1. *Nat Cell Biol* 2008;10:593–601.
- 27 Huang X, Ding L, Bennewith KL, *et al*. Hypoxia-inducible mir-210 regulates normoxic gene expression involved in tumor initiation. *Mol Cell* 2009;6:856–867.
- 28 Foekens JA, Sieuwerts AM, Smid M, *et al*. Four miRNAs associated with aggressiveness of lymph node-negative, estrogen receptor-positive human breast cancer. *Proc Natl Acad Sci USA* 2008;05:13021–13026.
- 29 Lawrie CH, Gal S, Dunlop HM, *et al*. Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. *Br J Haematol* 2008;5:672–675.
- 30 Wang J, Chen J, Chang P, *et al*. MicroRNAs in plasma of pancreatic ductal adenocarcinoma patients as novel blood-based biomarkers of disease. *Cancer Prev Res* 2009;9:807–813.
- 31 Patnaik SK, Kannisto E, Knudsen S, *et al*. Evaluation of microRNA expression profiles that may predict recurrence of localized stage I non-small cell lung cancer after surgical resection. *Cancer Res* 2010;70:36–45.
- 32 Alberg AJ, Nonemaker J. Who is at high risk for lung cancer? Population-level and individual-level perspectives. *Semin Respir Crit Care Med* 2008;29:223–232.
- 33 Seike M, Goto A, Okano T, *et al*. MiR-21 is an EGFR-regulated anti-apoptotic factor in lung cancer in never-smokers. *Proc Natl Acad Sci USA* 2009;106:12085–12090.
- 34 Brody JS, Spira A. State of the art. Chronic obstructive pulmonary disease, inflammation, and lung cancer. *Proc Am Thorac Soc* 2006;6:535–537.
- 35 Silverman EK, Spira A, Paré PD. Genetics and genomics of chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2009;6:539–542.
- 36 Steiling K, Lenburg ME, Spira A. Airway gene expression in chronic obstructive pulmonary disease. *Proc Am Thorac Soc* 2009;8:697–700.

Supplementary Information accompanies the paper on Modern Pathology website (<http://www.nature.com/modpathol>)