miR-1254 and miR-574-5p

Serum-Based microRNA Biomarkers for Early-Stage Non-small Cell Lung Cancer

Kristen M. Foss, BS, Chao Sima, PhD,* Donatella Ugolini, BS,†‡ Monica Neri, PhD,§ Kristi E. Allen, BS,* and Glen J. Weiss, MD*‖*

Introduction: The ability to diagnose non-small cell lung cancer (NSCLC) at an early stage may lead to improved survival. The aim of this study was to identify differentially expressed serum-based microRNAs (miRNAs) between patients with early-stage NSCLC and controls. These miRNAs may serve as biomarkers for NSCLC early detection.

Methods: miRNA profiling was performed on total RNA extracted from serum obtained from 22 individuals (11 controls and 11 patients with early-stage NSCLC). Quantitative polymerase chain reaction (qPCR) was used to validate the profiling results in the discovery set and in a validation set of 31 controls and 22 patients with early-stage NSCLC. Additionally, six matched plasma samples (four NSCLC cases and two controls) and three serum mesothelioma samples were analyzed by qPCR. Receiver operating characteristic curves were generated for each possible combination of the miRNAs measured by qPCR.

Results: The expression of hsa-miR-1254 and hsa-miR-574-5p was significantly increased in the early-stage NSCLC samples with respect to the controls. Receiver operating characteristic curves plotting these two miRNAs were able to discriminate early-stage NSCLC samples from controls with 82% and 77% of sensitivity and specificity, respectively, in the discovery cohort and with 73% and 71% of sensitivity and specificity, respectively, in the validation cohort. The mesothelioma and plasma samples did not seem to classify into either NSCLC or control groups.

Conclusions: Serum miRNAs are differentially expressed between patients with early-stage NSCLC and controls. The utility of miR-1254 and miR-574-5p serum-based biomarkers as minimally invasive screening and triage tools for subsequent diagnostic evaluation warrants additional validation.

Key Words: MicroRNAs, Biomarker, Early-stage disease, Non-small cell lung cancer, Early detection.

(*J Thorac Oncol.* 2011;6: 482–488)

Lung cancer is the most common cause of cancer-related death in the world. In the United States, more than 90 million individuals are at risk for developing lung cancer, and this disease is estimated to remain a major health problem for at least the next 50 years. More than 75% of lung cancer cases are diagnosed in late stages because there remains no practical way to screen a large number of people at risk. This is the major contributing factor to the dismal prognosis in non-small cell lung cancer (NSCLC), which accounts for 85% of lung cancers.[1] These results are clearly disappointing and point to the urgent need for new strategies to help identify those at-risk individuals that should be evaluated for this disease. Early detection offers the promise of improved cure rates. The dilemma remains as to how to efficiently facilitate stratification to appropriately identify these individuals at high risk for developing NSCLC. For NSCLC, broad application of computed tomography (CT) screening to at-risk populations, based on clinical parameters alone, has several drawbacks: (a) detection of significant rate of benign lung nodules at a rate up to 50%[2,3]; (b) cost-benefit analysis has projected that for every lung cancer death prevented by CT screening, approximately two false-positive invasive procedures will result[4]; (c) supply of adequately trained radiologists; and (d) coordination of appropriate follow-up and specialist referrals to investigate CT-detected abnormalities. Review of recent CT screening studies shows that although lung cancer diagnosis has increased more than threefold, the likelihood for thoracic resection for lung cancer has increased 10-fold.[5] Furthermore, there seems to be no meaningful reduction in the number of advanced cancers or the number of patients who die of lung cancer. This raises concerns for the use of CT for wide scale screening. Therefore, the need to find a less invasive, more reliable biomarker is crucial to increase the probability of earlier NSCLC detection and lower the risk of potential harm to

*Translational Genomics Research Institute, Phoenix, Arizona; †Department of Oncology, Biology and Genetics, University of Genoa, Genoa; ‡Unit of Epidemiology, Biostatistics and Clinical Trials, National Cancer Research Institute, Genoa; §Rehabilitative Pneumology, IRCCS San Raffaele Pisana, Rome, Italy; and ‖Virginia G. Piper Cancer Center at Scottsdale Healthcare, Scottsdale, Arizona.

Disclosure: Kristen Foss, Chao Sima, and Kristi Allen received grant funding from the Flinn Foundation; Donatella Ugolini and Monica Neri received grant funding from Associazione Italiana per la Ricerca Sul Cancro and Fondazione Buzzi ONLUS; and Glen Weiss received grant funding from the Flinn Foundation and has patent filings on the use of microRNA as theranostics.

Address for correspondence: Glen J. Weiss, MD, Scottsdale Clinical Research Institute, 10510 N 92nd St. Ste 200, Scottsdale, AZ 85258. E-mail: gweiss@tgen.org

ISSN: 1556-0864/11/0603-0482

TABLE 1. Clinical and Demographic Characteristics of Subjects in Discovery and Validation Cohorts

	Discovery Cohort		Validation Cohort		
	Control	NSCLC	Control	NSCLC	Mesothelioma
No. of cases	11[a]	11	31[b]	22	3
Median age, yr (range)	64 (60–74)	65 (50–72)	62 (39–91)	62 (52–77)	69 (62–76)
Gender					
Male	11	10	25	19	2
Female	0	1	6	3	1
Smoking history					
Never smoker	1	0	0	1	1
Current smoker	5	6	11	6	1
Former smoker	5	5	20	15	1
Median pack years (range)	23.5 (0–46.5)[c]	42 (4.35–71.25)	44 (18–65)	42.5 (0–112.5)[d]	38 (0–46)
Stage					
I	NA	4	NA	6	1
I/II	NA	0	NA	10[e]	0
II	NA	6	NA	4	2
Unavailable	NA	1	NA	2	0
Histology					
Adenocarcinoma	NA	6	NA	10	NA
Bronchioloalveolar carcinoma	NA	1	NA	2	NA
Squamous cell carcinoma	NA	1	NA	6	NA
Large cell carcinoma	NA	1	NA	2	NA
Other	NA	1	NA	1	NA
Unavailable	NA	1	NA	1	NA

[a] Two cases have two replicates that were treated as separate samples in the analysis.
[b] One case was diagnosed with bladder cancer at the end of 2009, 1.5 yr after sample collection.
[c] Subject with 0 pack years has smoked four cigars a day for 30 yr.
[d] Missing data for two cases.
[e] T2N0, no subclassification for T2a or T2b available. Therefore, cannot distinguish whether stage I or II.
NA, not applicable; NSCLC, non-small cell lung cancer.

truly make a difference in survival and at the same time, remain both cost and resource utilization effective.

Blood serum is a minimally invasive, low risk, and easy to obtain biofluid that can also be explored for potential biomarkers. Serum-based microRNAs (miRNAs) have been shown to be differentially expressed in various patients with cancer, including NSCLC.[6] Disease characterization has also been demonstrated using miRNA expression profiling.[7] These small noncoding RNAs of 21 to 25 nucleotides are ideal for expression profiling compared with larger messenger RNAs that are more susceptible to degradation[6] and often poorly overlap among various datasets.[8]

The aim of this study was to identify serum-based miRNA biomarkers that discriminate early-stage NSCLC from healthy individuals with a tobacco smoking history by quantitative reverse transcription polymerase chain reaction (qRT-PCR). The ability to identify patients with NSCLC through a minimally invasive, potentially cost-effective manner could lead to earlier detection of NSCLC and potentially improved outcomes for these patients.

PATIENTS AND METHODS

Patient Samples

Serum and plasma samples from patients with (1) early-stage NSCLC or (2) mesothelioma and (3) healthy individuals (controls) with a history of smoking tobacco were used in this study. Note: all samples collected from cancer cases were done before initiation of cancer-directed therapy. The discovery cohort consisted of serum samples from 11 controls and 11 patients with NSCLC. The validation cohort consisted of serum samples taken from 22 patients with NSCLC and 31 controls. Additionally, serum from three early-stage mesothelioma patients and a total of six plasma samples from both discovery and validation cohorts (four NSCLC and two controls) were also available for analysis (Table 1). All samples were acquired from the Cancer of RESpiratory Tract biorepository under a standard protocol described elsewhere.[9] Appropriate informed consent was obtained and institutional review board approvals were in place.

RNA Extraction and Cel-miRNA Spike in

Total RNA was isolated from a volume of 75 to 200 μl of serum or plasma using phenol and guanidine thiocyanate. A total of 2 ng of cel-miR-39 was spiked into each sample[10,11] after the addition of the phenol and guanidine thiocyanate to serve as an external processing control.

miRNA Microarray Profiling

A minimum of 100 ng of total RNA from the discovery cohort was added to the GenoExplorer microRNA Expression System (GenoSensor Corporation, Tempe, AZ) containing

probes in triplicate for 880 mature miRNAs (Sanger miRNA Registry version 13, September 2008) and 473 premiRNAs along with positive and negative control probes. SAM data analysis was applied to find significantly differentially expressed miRNAs in one condition in contrast to the other. Data were normalized to PC-U6B, U6-337, 5S-rRNA, and PC-HU5S. The top differentially expressed mature miRNAs in serum from patients with NSCLC versus controls were selected based primarily on fold change more than 1.54, followed by *p* values, *q* values, and false discovery rates.

qRT-PCR Analysis of miRNAs

qRT-PCR was performed on the selected miRNA candidates to validate the miRNA array results. The GenoExplorer miRNA First-Strand cDNA Core Kit (2002-50, GenoSensor Corporation) was used to generate miRNA first-strand complementary DNA. miRNA expression levels were measured using SYBR Green (04887352001; Roche; Indianapolis, IN). miRNA-specific forward primers and a universal reverse primer were purchased (GenoSensor Corporation). The reaction conditions were 15 minutes of denaturation at 94°C followed by 45 cycles of 94°C for 30 seconds, 59°C for 15 seconds, and 72°C for 30 seconds. Melting curve analysis was used to assess the specificity of the amplified product. All qRT-PCR reactions were carried out in triplicate on the Lightcycler 480 (Roche). miRNA expression was normalized to the expression of RNU6 and cel-miR-39 separately. We normalized using a variation of the "delta" method. Specifically, if the Cp values for a miRNA and the normalizer are C1 and C0, respectively, the normalized value for C1 is $1 - (C1 - C0)/\text{maxCp}$, where maxCp is the number of maximum cycles.

qRT-PCR Data Analysis

A logit regression model was fit on the PCR data with the predictors being one or more miRNAs (a predictor set, therefore, comprises different combination of miRNAs). For each logit classifier using a predictor set, a receiver operating characteristic (ROC) curve was plotted, and area under the curve (AUC) was computed. The best sensitivity/specificity pair is selected as the highest specificity with sensitivity being at least 0.6. *p* values for the predictor set using *t* statistical test and Mann–Whitney–Wilcoxon test are also computed using coefficients obtained from logit regression.

RESULTS

RNA Extracted from Human Samples

The 24 available samples from the discovery cohort had a median of 352 ng total RNA (range: = 147–1762 ng) derived from 200 μl serum and proceeded to miRNA microarray profiling and subsequent qRT-PCR analysis. Baseline characteristics for the discovery cohort are listed in Table 1. The median age was 64 and 65 years for controls and NSCLC, respectively. Nearly all were men, and all but one individual (control no. 2678) had a history of tobacco smoking in both groups. The median follow-up time for the discovery cohort controls was 33.4 months (range: 27.4–114.9 months). The majority of the discovery cohort NSCLC cases were adenocarcinoma. Samples for the validation cohort became available for analysis at a later date. The 53 serum samples from the validation cohort had a median of 870 ng total RNA (range = 410–3796 ng) derived from 100 to 200 μl serum and proceeded to qRT-PCR analysis. Baseline characteristics for the validation cohort are listed in Table 1. In both controls and NSCLC, the median age was 62 years, with a majority of men, and nearly all had a documented history of tobacco smoking. The median follow-up time for the validation controls was 40.9 months (range: 8.8–108.5 months). Adenocarcinoma was the most common NSCLC histology in the validation cohort. Total RNA was also extracted from the three serum samples from early-stage mesothelioma patients and six plasma samples from NSCLC and control for qRT-PCR analysis. Detailed baseline clinical characteristics for all samples are shown in Supplementary Table 1 (http://links.lww.com/JTO/A57).

miRNA Profiling Analysis Reveals Differentially Expressed miRNAs

miRNA profiling was performed on 24 samples of the discovery cohort (including two separate serum samples from control no. 1075 and control no. 1462). There were eight miRNAs significantly (p value <0.05) differentially expressed by microarray, and four of the top 25 mature miRNAs had fold change more than 1.54. Because none of the top hits had false discovery rate less than 0.05, we proceeded to test the four mature miRNAs with fold change more than 1.54 between the patients with NSCLC and the controls to allow for greatest distribution to stratify between the two groups when proceeding to qRT-PCR measurement (Table 2). None of the candidate miRNAs were significantly down-regulated in early-stage NSCLC. To avoid overfitting, we intentionally did not consider a larger signature due to the sample size.

Validation of miRNA Array Findings by qRT-PCR in Discovery Cohort

The four miRNAs identified by array profiling were measured by qRT-PCR in the discovery cohort. All possible

TABLE 2. miRNA Array Results

miRNA	Ratio LCS/CS Median	*p*	FDR	*q* Value
Hsa-miR-1268	2.22	0.06	0.28	0.12
Hsa-miR-574-5p	2.17	0.04	0.22	0.12
Hsa-miR-1254	1.62	0.21	0.42	0.12
Hsa-miR-1228*	1.54	0.20	0.45	0.12
Hsa-miR-297	1.43	0.03	0.30	0.12
Hsa-miR-1225-5p	1.39	0.01	0.31	0.12
Hsa-miR-923	1.23	0.03	0.24	0.12
Hsa-miR-1275	1.23	0.04	0.25	0.12
Hsa-miR-185	1.16	0.01	0.57	0.12
Hsa-miR-483-5p	1.13	0.03	0.26	0.12
Hsa-miR-320a	1.10	0.02	0.36	0.12

Median signal intensity normalized to the expression of PC-U6B, U6-337, 5S-rRNA, and PC-HU5S was used for analysis. miRNAs in bold were selected for qPCR validation.

LCS, lung cancer serum; miRNA, microRNA; CS, control serum; FDR, false discovery rate; qPCR, quantitative polymerase chain reaction.

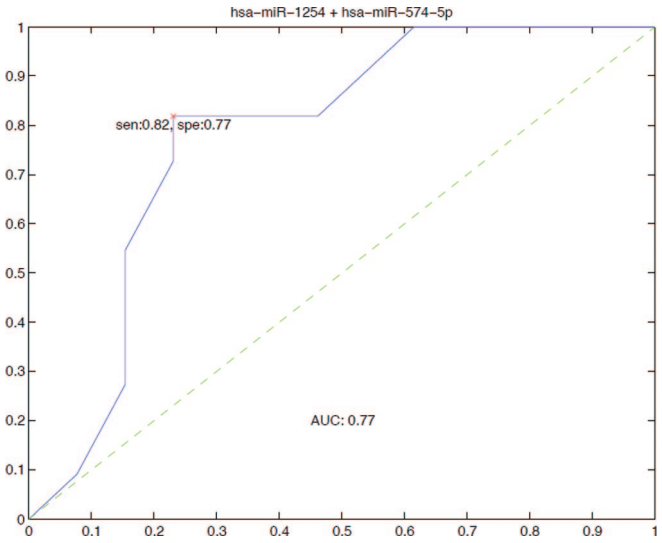

FIGURE 1. miR-1254 and miR-574-5p receiver operating characteristic (ROC) curve for discovery cohort. Quantitative polymerase chain reaction (qPCR) expression of miR-1254 and miR-574-5p was best able to stratify patients with early-stage non-small cell lung cancer (NSCLC) from controls with an area under the curve (AUC) of 0.77 and a 82% and 77% of sensitivity and specificity, respectively. Data were normalized to expression of cel-miR-39.

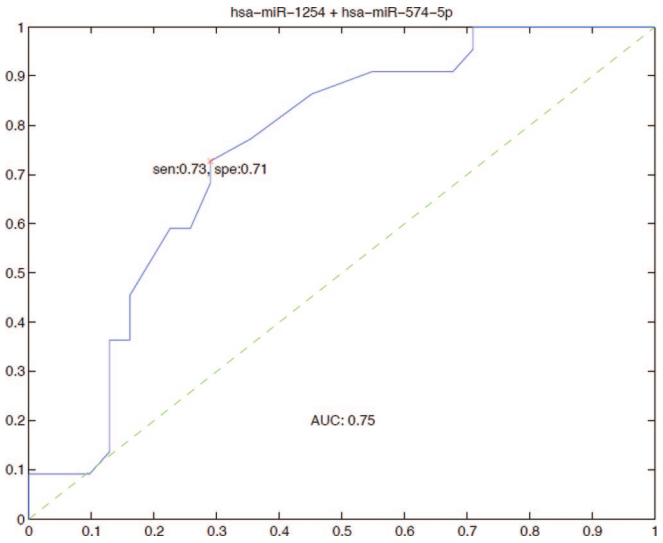

FIGURE 2. miR-1254 and miR-574-5p receiver operating characteristic (ROC) curve for validation cohort. Curves were generated from quantitative polymerase chain reaction (qPCR) microRNA (miRNA) data normalized to cel-miR-39. Stratification of early-stage non-small cell lung cancer (NSCLC) and controls was achieved with a sensitivity and specificity of 73% and 71%, respectively. The area under the curve (AUC) is 0.75, and the model is correctly classified for 16 of 22 NSCLC cases and 22 of 31 control cases.

combinations of these four miRNAs were plotted on ROC curves to determine the combination that would stratify the patients with the highest sensitivity and specificity. The best set of predictors included two miRNAs, hsa-miR-1254 (miR-1254) and hsa-miR-574-5p (miR-574-5p), that were differentially expressed between the controls and stage I to stage II NSCLC patient samples (p values of 0.016 and 0.0277 with t statistical test and Mann–Whitney–Wilcoxon test, respectively) with an AUC of 0.77 and can achieve 82% and 77% for sensitivity and specificity, respectively (Figure 1). This data suggest the existence of significantly differentially expressed serum-based miRNAs between patients with NSCLC and controls.

Measuring miRNA Biomarkers in a Validation Cohort

Expression of miR-1254 and miR-574-5p was measured in the serum from 53 NSCLC and control cases. When normalized to cel-miR-39, the ROC curve had an AUC of 0.75 and a 73% and 71% of sensitivity and specificity, respectively (p values of 0.0123 and 0.0025 with t statistical test and Mann–Whitney–Wilcoxon test, respectively) (Figure 2). Note: the results were normalized separately to RNU6 and cel-miR-39, and the ROC curves show little difference between normalizing to RNU6 and normalizing to cel-miR-39 for both the discovery and validation cohorts with coefficient of variation less than 10% (data not shown).

No differences were seen in the sensitivity or specificity when applying miR-1254 and miR-574-5p classifiers to adenocarcinoma or gender-specific analysis (data not shown). Expression of miR-1254 and miR-574-5p was also measured in the serum of three early-stage mesothelioma cases. The available plasma from four NSCLC and two controls did not consistently classify into either the serum NSCLC or the serum control group (data not shown).

DISCUSSION

The main result of this study is that miRNAs are differentially expressed between patients with early-stage NSCLC and controls. In particular, the expression of hsa-miR-1254 and hsa-miR-574-5p was significantly increased in early-stage NSCLC serum samples with respect to the controls and could be detected even months before clinical diagnosis. The combination of these two miRNAs may be a useful biomarker for early diagnosis and a handy tool when screening large numbers of high-risk individuals.

In addition, we confirm that miRNAs can be measured from a relatively small amount of serum.[10]

Approximately 186,000 new NSCLC cases are diagnosed annually in the United States, and approximately 140,000 will be diagnosed with late-stage disease. To make any dramatic impact to improve survival, identification of high-risk individuals (primarily those with tobacco smoking history) is sorely needed. Wide-based population screening using general clinical risk factors alone has several significant drawbacks including high-false positives, cost, supply of trained health care staff, follow-up, and associated procedural and psychological morbidities. Initial triage for image-based evaluation using other minimally and/or noninvasive testing may be the key to alleviating some of these drawbacks.

TABLE 3. Summary of Other Studies Investigating Lung Cancer Biomarkers

Authors	Chen et al.[6]		Patz et al.[16]		Young et al.[17]		Yu et al.[18]		Spira et al.[12]		Showe et al.[13]		Keller et al.[14]
Sample set	Initial	Test	Initial	Test	Initial	Test	Initial	Test	Initial	Test	Initial	Test	Initial
No. of samples	32	237	100	97	439	491	72	122	129	35	228	55	36
Control	21[a]	75	50	48	200	284	36	58	69	17	91	17	19
Lung cancer	11[a]	152	50	49	239	207	36	64	60	18	137	38	17
Independent validation sample set?[b]	Not accurately described		No		No		No		Yes		Yes[c]		No
Gender													
Control													
Male	11	Unavailable	28	23	290[d]		Unavailable	35	51	12	55	Unavailable	7
Female	10	Unavailable	22	25	194[d]		Unavailable	23	18	5	36	Unavailable	12
Lung cancer													
Male	Unavailable		29	27	236		Unavailable	39	48	12	69	Unavailable	9
Female	Unavailable		21	22	210		Unavailable	25	12	6	68	Unavailable	8
No. of stage I and II cases	Unavailable		22	19	NA		36	31	13	NA	93	31	13
Histology													
Non-small cell	Unavailable		50	45	Unavailable		36	64	48	Unavailable	137	38	17
Adenocarcinoma	Unavailable		16	22	Unavailable		36	33	10	Unavailable	85	29	7
Bronchioloalveolar carcinoma	Unavailable		0	4	Unavailable		0	0	1	Unavailable	0	0	1
Squamous cell carcinoma	Unavailable		13	9	Unavailable		0	31	23	Unavailable	42	7	7
Large cell carcinoma	Unavailable		1	0	Unavailable		0	0	4	Unavailable	0	0	0
Other	Unavailable		20	10	Unavailable		0	0	10	Unavailable	10	2	2
Small cell	Unavailable		0	4	Unavailable		0	0	11	Unavailable	0	0	0
Unavailable	Unavailable		0	0	Unavailable		0	0	1	Unavailable	0	0	0
Sample type	Serum		Serum		Blood		Sputum		Large airway epithelial cells		Peripheral blood mononuclear cells		Blood
Type of biomarker	microRNA		Protein		SNPs		microRNA		Gene expression		Gene expression		microRNA
No. of biomarkers in classifier	2		6		20		4		80		29		24
Sensitivity/specificity of classifier (%)	NA		89.3/84.7	77.8/75.4	Unavailable		80.6/91.7	70.3/80	80/84	83/76	91/80	76/82	92.5/98.1

[a] Samples were pooled before RNA extraction.
[b] Independent validation sample set defined as test set samples collected from a different biorepository than initial sample set.
[c] Independent validation set contained a mix of samples from the same biorepository and another collection protocol within the institution.
[d] No breakdown between initial and test set was reported.
NA, not applicable; SNP, single-nucleotide polymorphism.

Several groups have explored other modalities for biomarker utility as a potential early-detection screening tool for NSCLC. We selected studies reporting evaluation of biomarkers for early detection of lung cancer and validation in an independent cohort or miRNA biomarkers for early detection of lung cancer (Table 3). All these studies have certain limitations including sample size and detailing the clinical follow-up time for evaluated control samples. Some of these studies included patients with stage III or IV NSCLC or even small cell lung cancer patient samples. Samples analyzed spanned from blood, peripheral blood mononuclear cells, sputum, serum, to large airway epithelial cells. The type of biomarkers explored included miRNA, messenger RNA, protein, or single-nucleotide polymorphisms. Limitation for most of these studies, including ours, is the lack of validation of the biomarker classifier in a sample set collected from a biorepository separate and independent from where the discovery samples were deposited. Another limitation for most of these studies is a lack of documentation of the follow-up time for samples collected from control subjects. As most controls tend to be clinically matched to the NSCLC cases, they also have the potential risk to develop NSCLC at a later date. Misclassification of a sample because of inadequate follow-up can potentially impact the utility of the biomarker classifier. All together, our control samples had a median follow-up time of 39.1 months, without developing NSCLC. The Cancer of Respiratory Tract biorepository is able to query the Italian tumor registry to identify potential misclassifications months to years after the sample has been banked. Samples banked months to years before an individual is diagnosed with early-stage NSCLC are extremely valuable to test the potential utility of a NSCLC biomarker as an early-detection tool. This lead time information is often lacking from early-detection biomarker studies. Two of the NSCLC cases in our study were collected more than 8 months before diagnosis, although true confirmation of early-stage disease is not available. We acknowledge that our sample size for discovery and validation is small, but compared with works of others listed in Table 3, the number of stage I and stage II NSCLC is similar or smaller to our sample size. It is important to validate our classifier in a larger sample set of stage I/II NSCLC and matched controls from an independent biorepository collection.

Only two studies in Table 3 reasonably describe the source of their validation samples as derived from an independent biorepository collection.[12,13] Nevertheless, both these studies fail to adequately describe the clinical characteristics of their validation set,[12,13] particularly the number of early-stage NSCLC.[12] This could impair the utility of applying these "meta-gene" classifiers to differentiate early-stage NSCLC from controls.

In a majority of the studies, the number of biomarkers proposed for further study total close to or more than the total number of samples analyzed, which remains a concern for data overfitting or attests to the minimal real contribution of each individual biomarker among the entire classifier. For example, Keller et al.[14] used a 24 "meta-miR" support vector machine classifier after trials of five different support vector machine kernel functions and multiple feature sizes and reported a sensitivity and specificity of 92.5% and 98.1%, respectively. It is yet to be seen how this classifier performs on an independent sample set.

The expression of miR-1254 is associated with early-stage cancer for the first time in this study, whereas miR-574-5p is listed as a differentially expressed miRNA in two previous lung cancer miRNA publications.[6,14] Our group has also recently reported on miR-574-5p's significant association with chemoresistance in small cell lung cancer.[15]

Serum samples were available from three stage I to stage II pleural malignant mesothelioma patients. Although limited in number, these samples are highly valuable given the rarity of this severe asbestos-linked cancer and the fact that in the large majority of cases, mesothelioma is diagnosed at late stages. Although very preliminary, our results suggest that miR-1254 and miR-574-5p could not classify mesothelioma to the same group as NSCLC. miRNA expression in mesothelioma deserves new dedicated studies, as there is a deep need of effective biomarkers to screen the cohorts of asbestos exposed subjects at risk for mesothelioma.

Parallel plasma samples were analyzed in some subjects, in addition to serum. Expression of miR-1254 and miR-574-5p in plasma did not seem to differentiate between early-stage NSCLC and controls. There was no difference in the way the serum and plasma samples were processed for RNA extraction or analyzed.

In conclusion, we propose the serum-based miR-1254 and miR-574-5p biomarkers for future application as minimally invasive screening and triage tools for subsequent diagnostic evaluation, after additional validation in samples collected by other biorepositories.

ACKNOWLEDGMENTS

Supported by the Flinn Foundation (to K.M.F., C.S., K.E.A., and G.J.W.), Associazione Italiana per la Ricerca Sul Cancro (AIRC), Italy; University of Genoa, Italy; Fondazione Buzzi ONLUS, Italy.

The authors are grateful to the patients who contributed samples for this project. They thank the Flinn Foundation for supporting this project and Russ Brandt for logistical support on this project. They acknowledge James Xia, Shravan Sridhar, David Edwards V, and Carlos Lorenzo for assistance on this project.

REFERENCES

1. Jemal A, Siegel R, Ward E, et al. Cancer statistics, 2009. *CA Cancer J Clin* 2009;59:225–249.
2. Swensen SJ, Jett JR, Hartman TE, et al. CT screening for lung cancer: five-year prospective experience. *Radiology* 2005;235:259–265.
3. Henschke CI, Yankelevitz DF, Libby DM, et al. Early lung cancer action project: annual screening using single slice helical CT. *Ann NY Acad Sci* 2001;952:124–134.
4. Mahadevia PJ, Fleisher LA, Frick KD, et al. Lung cancer screening with helical computed tomography in older adult smokers: a decision and cost effectiveness analysis. *JAMA* 2003;289:313–322.
5. Bach PB, Jett JR, Pastorino U, et al. Computed tomography screening and lung cancer outcomes. *JAMA* 2007;297:953–961.
6. Chen X, Ba Y, Ma L, et al. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* 2008;1–10.

7. Rosenfeld N, Aharonov R, Meiri E, et al. MicroRNAs accurately identify cancer tissue origin. *Nat Biotechnol* 2008;26:462–469.
8. Lau SK, Boutros PC, Pintilie M, et al. Three-gene prognostic classifier for early-stage non small-cell lung cancer. *J Clin Oncol* 2007; 25:5562–5569.
9. Ugolini D, Neri M, Canessa PA, et al. The CREST biorepository: a tool for molecular epidemiology and translational studies on malignant mesothelioma, lung cancer, and other respiratory tract diseases. *Cancer Epidemiol Biomarkers Prev* 2008;17:3013–3019.
10. Kroh EM, Parkin RK, Mitchell PS, et al. Analysis of circulating microRNA biomarkers in plasma and serum using quantitative reverse transcription-PCR (qRT-PCR). *Methods* 2010;50:298–301.
11. Mitchell PS, Parkin RK, Kroh EM, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci USA* 2008;105:10513–10518.
12. Spira A, Beane JE, Shah V, et al. Airway epithelial gene expression in the diagnostic evaluation of smokers with suspect lung cancer. *Nat Med* 2007;13:361–366.
13. Showe MK, Vachani A, Kossenkov AV, et al. Gene expression profiles in peripheral blood mononuclear cells can distinguish patients with non-small cell lung cancer from patients with nonmalignant lung disease. *Cancer Res* 2009;69:9202–9210.
14. Keller A, Leidinger P, Borries A, et al. miRNAs in lung cancer—studying complex finger prints in patient's blood cells by microarray experiments. *BMC Cancer* 2009;9:353.
15. Ranade AR, Cherba D, Sridhar S, et al. MicroRNA 92a-2*, a biomarker predictive for chemoresistance and prognostic for survival in small cell lung cancer patients. *J Thoracic Oncol* 2010;5:1273–1278.
16. Patz EF, Campa MJ, Gottlin EB, et al. Panel of serum biomarkers for the diagnosis of lung cancer. *J Clin Oncol* 2007;25:5578–5583.
17. Young RP, Hopkins RJ, Hay BA, et al. A gene-based risk score for lung cancer susceptibility in smokers and ex-smokers. *Postgrad Med J* 2009;85:515–524.
18. Yu L, Todd NW, Xing L, et al. Early detection of lung adenocarcinoma in sputum by a panel of microRNA markers. *Int J Cancer* 2010;127: 2870–2878.