ORIGINAL RESEARCH

Wilson H. Roa, MD¹
Julian O. Kim, MD¹
Rene Razzak, MD^{1,2}
Hongfei Du, MD³
Linghong Guo, PhD¹
Ravinder Singh, PhD¹
Sayf Gazala, MD²
Sunita Ghosh, PhD¹
Eric Wong, MD⁴
Anil A. Joy, MD¹
James Z. Xing, MD⁵
Eric L. Bedard, MD⁶

¹ Department of Oncology, Cross Cancer Institute and University of Alberta, Edmonton, Alberta, Canada,

² Division of General Surgery, Department of Surgery, University of Alberta, Edmonton, Alberta, Canada,

³ School of Mathematical Sciences, University of Electronic Science and Technology of China, Chengdu, China,

⁴ Division of Pulmonary Medicine, Department of Medicine, University of Alberta Hospital & University of Alberta, Edmonton, Alberta, Canada,

⁵ Department of Laboratory Medicine and Pathology, University of Alberta, Edmonton, Alberta Canada,

⁶ Division of Thoracic Surgery, Department of Surgery, Royal Alexandra Hospital and University of Alberta, Edmonton, Alberta, Canada

Manuscript submitted 20th January, 2012.

Manuscript accepted 4th September, 2012.

Clin Invest Med 2012; 35 (5): E271-E281.

Sputum MicroRNA Profiling: A Novel Approach for the Early Detection of Non-Small Cell Lung Cancer

Abstract

Purpose: MicroRNAs (miRNAs) post-transcriptionally regulate hundreds of gene targets involved in tumorigenesis thereby controlling vital biological processes, including cellular proliferation, differentiation and apoptosis. MiRNA profiling is an emerging tool for the potential early detection of a variety of malignancies. This study was conducted to assess the feasibility and methodological robustness of quantifying sputum miRNAs, employing quantitative real-time polymerase chain reaction (RT-qPCR) and cluster analysis on an optimized miRNA profile as a novel approach for the early detection of non-small cell lung cancer (NSCLC).

Methods: The relative expressions of 11 miRNAs in sputum (miR-21, miR-145, miR-155, miR-205, miR-210, miR-92, miR-17-5p, miR-143, miR-182, miR-372, and let-7a) in addition to U6 were retrospectively assessed in four NSCLC-positive and four negative controls. Subsequently, a set of five miRNAs (miR-21, miR-143, miR-155, miR-210, miR-372) was selected because of degree of relatedness observed in the cluster analysis and tested in the same sputum sample set. The five optimized miRNAs accurately clustered these eight retrospective patients into NSCLC positive cases and negative controls. The five miRNA panel was then prospectively quantified in the sputum of 30 study patients (24 NSCLC cases and six negative controls) in a double-blind fashion to validate a five miRNA panel using hierarchical cluster analysis.

Results: The optimized five miRNA panel detected NSCLC (83.3% sensitivity and 100% specificity) in 30 prospectively accrued study patients.

Conclusion: Sputum miRNA profiling using cluster analysis is a promising approach for the early detection of non-small cell lung cancer. Further investigation using this approach is warranted.

Correspondence to:

Dr. Wilson Roa
Department of Radiation Oncology
Cross Cancer Institute
11560 University Avenue
Edmonton, Alberta, Canada T6G 1Z2
Email: Wilson.Roa@albertahealthservices.ca

MicroRNAs (miRNAs) are a group of recently discovered small non-protein-coding RNAs. MiRNAs post-transcriptionally regulate the expression of hundreds of target genes, thereby controlling a wide range of tumorigenic processes including cellular proliferation, differentiation, and apoptosis [14]. MiRNAs are expressed in a tissue-specific manner and are, therefore, candidate biomarkers for cancer detection and non-malignant disease processes [15-20]. Furthermore, miRNAs may function as tumor suppressors or tumor promoters (oncomirs), with dysregulated miRNA expression leading to carcinogenesis. [9, 16, 21-22]. MiRNAs have shown potential for the diagnosis and classification of human malignancies [14-16]. MiRNA dysregulation (over- and under-expression) in surgically resected lung tumor tissues is reliably correlated with the diagnosis and prognosis of the lung cancer patients [14-16, 21-22]. It is becoming evident that miRNA-based detection techniques could improve upon cytology (with its low sensitivity) and of computed tomography (with its high false-positive rate), and seem to hold promise to develop screening tools for the early detection and monitoring of lung cancers.

There are few available studies that have examined the feasibility of aberrant miRNA expression profiles in sputum as a potential screening tool for lung cancer. In one such study, Xie et al. reported that miR-21 over-expression can be used for the diagnosis of NSCLC with a sensitivity of 69.6% [9]; however, a single miRNA lacks sensitivity and specificity as a cancer biomarker because lung cancer is a heterogeneous entity, with oncogenesis arising through a poorly understood, complex, multi-step process [13, 15, 22-26]. There is considerable interest in using a combination of specific miRNAs (a miRNA profile), for use in the early detection of lung cancer [24-26] in patients. Yu et al. reported that four miRNAs (miR-21, miR-486, miR-375 and miR-200b) distinguished lung adenocarcinoma patients from normal subjects with a sensitivity and specificity of 80.6% and a 91.7%, respectively [24]. Xing et al. showed that the profile of miR-205, miR-210, miR-708 detected NSCLC with a relatively low sensitivity (73%). [25]. It is apparent that the selection and optimization of miRNAs and miRNA profile quantitative analysis methodology requires further investigation prior to its widespread acceptance and implementation as a screening tool [27-29].

Cluster analysis is an analytical mathematical method that identifies groups of samples that perform similarly or possess similar characteristics. Cluster analysis can lead to a diagnosis when a distinct grouping, based on the data, is observed [30]. In this study, hierarchical cluster analysis, a special clustering algorithm that is based on a tree-like structure, was used to study the relationship between the observations of miRNA

profiles to differentiate patients with NSCLC from healthy controls [31].

The objective of this study was to determine the optimum sputum miRNA profile for use as a future screening test for early NSCLC detection. Eleven miRNAs that have been shown to be dysregulated, and have variable expression strongly correlated with the presence or absence of cancers and NSCLC, were selected [13-27]. An optimized five miRNA panel was then determined retrospectively on a small subset of eight study patients and further validated prospectively in a larger subset of patients for the diagnosis and early detection of NSCLC.

Materials and Methods

Patient selection

This study was approved by the Human Research Ethics Review Board of the University of Alberta (Edmonton, Alberta, Canada), as well as the Alberta Cancer Research Ethics Committee (Alberta Health Services, Edmonton, Alberta, Canada). Written informed consents were obtained from all participating patients prior to study entry. Patients with potentially resectable, pathologically-confirmed NSCLC were considered eligible for inclusion regardless of stage. For negative controls, eligible patients were those with no current or previous history of malignancy and recent chest imaging (Chest X-Ray or CT-Thorax within the past 12 months) that demonstrated no evidence of pulmonary malignancy. Negative controls were not excluded if they had active or previous history of non-malignant pulmonary disease (i.e., asthma, emphysema and interstitial lung disease) or if they smoked.

Collection and treatment of sputum

Prior to the collection of a sputum sample, patients rinsed their mouths with water, breathed deeply, held their breath and coughed. All expectorated sputum was collected into a sterile plastic sample container that was then sealed and stored at -20°C until further processing. Since previous studies have shown that sputum cytology has a high rate of both false positives and false negatives [38], routine sputum cytology was not performed on the provided sputum samples. Sputum samples were homogenized and 200 µl of the sputum samples were transferred to individual nuclease-free tube (1.5 ml). In order to fully homogenize the sputum, 400 µl of sputolysin solution (Sigma Aldrich, USA) was added and the samples vortexed at 3000 rpm for 30 seconds then incubation at 37°C for 30 minutes. Homogenized sputum samples were subsequently stored at -20 oC until RNA extraction was performed.

RNA isolation

TRIzol (1.0 ml) (Invitrogen, USA) was added to the individual homogenized sputum samples. The samples were then vortexed at 3,000 rpm for 20 seconds and reacted at room temperature for 5 minutes. Trichloromethane (200 μ l) (Sigma-Aldrich, USA) was added to extract RNA then 500 μ l of isopropyl alcohol was added to precipitate RNA, which was then washed with 75% EtOH. The RNA was then dissolved in nuclease-free water. The total RNA was quantified using a UV-spectrometer (Du-7000, Beckman, USA).

RNA reverse transcription

Based on a thorough evaluation of the best available literature at the time of study design, 11 miRNAs and one endogenous control (U6) were selected for evaluation. The 11 miRNAs investigated were as follows: miR-21, miR-145, miR-155, miR-205, miR-210, miR-92, miR-17-5p, miR-143, miR-182, miR-372, let-7a. The TaqMan MicroRNA Reverse Transcription Kit for individual miRNAs (Applied Biosystems, USA) was used for reverse transcription (RT). The RT reaction mixture comprised of 50 nM stem-loop miRNA-specific primer (Applied Biosystems, USA), 1X RT buffer (Applied Biosystems, USA), 0.25 mM each of dNTPs (Applied Biosystems, USA), 3.33 U/ μ l MultiScribe reverse transcriptase (Applied Biosystems, USA), and 5 μ l RNA (~0.1 μ g RNA) in a total volume of 15 μ l. The RT was performed using StepOnePlus RT-PCR instrument (Applied Biosystem, USA). The RT products were stored at -20 oC until subsequent RT-PCR analysis.

Quantitative real-time Polymerase Chain Reaction (RT-qPCR)

The RT-qPCR (TaqMan) miRNA assays (Applied Biosystems, USA) for individual miRNAs were performed in duplicate using the RT reaction derived from a single sputum sample from each patient. The RT-qPCR was carried out using a StepOnePlus RT-PCR instrument (Applied Biosystems, USA). Each PCR reaction included 2 μ l RT product, 10 μ l of 1X TaqMan Universal PCR Master Mix, 1 μ l stem-loop miRNA-specific primer and probe [28] in 20 μ l of final reaction volume. The PCR reaction contents were incubated in a 96-well plate at 95°C for 15 minutes followed by 45 cycles of 95°C for 15 seconds and 60°C for 1 minute. The threshold cycle (CT) was defined as the fractional cycle number at which the fluorescence passed the fixed threshold. SDS software (Applied Biosystems, USA) was used to automatically identify cycle threshold (CT) values. The expression levels of the 11 individual target miRNAs were normalized to U6, a commonly used inter-

nal control for miRNA quantification assays, and a reference sample consisting of normal lung fibroblast (MRC-5) [32-34]. The comparative method ($\Delta\Delta C_T$ method) was used to quantify the RT-qPCR data for miRNA expression. MiRNA expression data is presented as the fold change in miRNA expression normalized to the endogenous control (U6) and relative to the MRC-5 reference sample. Using $\Delta\Delta C_T$ method, the relative miRNA expression in testing sample is expressed as:

$$RN=2^{-\Delta\Delta C_T}$$

$$\Delta\Delta C_T = (C_{T,m} - C_{T,ec})_{\text{sample}} - (C_{T,m} - C_{T,ec})_{\text{reference}}$$

Where: $C_{T,m}$ is the C_T for the miRNA measured and $C_{T,ec}$ is the C_T for the endogenous control miRNA (U6) for the sample to be tested (sputum) and the reference sample (MRC-5 normal lung cell line); RN is the amount of target miRNA required to be tested.

MiRNA panel optimization

In order to optimize the miRNA panel, the 11 miRNA panel was retrospectively applied to the sputum samples of four known NSCLC-positive cases and four known NSCLC-negative controls. The sample collection and RT-qPCR analysis were performed in a double-blinded manner. Cluster analysis of the 11 miRNA variables was then performed. MiRNAs that were observed to exhibit a greater degree of relatedness were chosen to form a more selective and optimized panel. This selective panel was then applied to the expression profiles of the eight patients to confirm the ability to successfully cluster the NSCLC-positive cases from the negative controls.

Prospective miRNA validation

The miRNA expression profiles from the sputum of 30 prospectively accrued patients from a thoracic surgery clinic (24 NSCLC positive cases and six negative controls) were examined. Hierarchical cluster analysis was performed on the relative miRNA expressions to validate both our selective miRNA panel and analytical methodology used in the experiments. Sample collection and analysis was carried out in a double-blind fashion.

Statistical analysis

Variable miRNA expression provides each individual's sputum sample with a distinctive pattern of expression. The experimental-normalized miRNA expression profiles were used to evaluate the sputum samples using average linkage and correlation similarity. The nearest-neighbour cluster method and squared Euclidean distance were used to measure the intervals. Cluster analysis was performed using the hierarchical cluster

TABLE 1. Clinical and pathological characteristics of the study cohort.

Study ID	NSCLC Status (+ or -)	Histological Subtype	Stage (AJCC7 th ed)	Age	Sex	Smoking history
Prospectively Accrued Patients						
N01	+	Adenocarcinoma	IB	71	F	Ex-smoker
N02	+	Adenocarcinoma	IA	63	M	Current smoker
N05	+	Adenocarcinoma	IA	64	M	Current smoker
N06	+	Adenocarcinoma	IIB	65	M	Current smoker
N07	+	NSCLC NOS	IB	80	F	Current smoker
N08	+	Adenocarcinoma	IA	71	F	Ex-smoker
N10	+	Squamous Cell Ca.	IA	80	M	Current smoker
N12*	+	NSCLC NOS	IA	63	M	Current smoker
N14	+	Squamous Cell Ca.	IA	66	M	Current smoker
N16	+	Adenocarcinoma	IIB	72	M	Ex-smoker
N18*	+	Adenocarcinoma	IIA	51	M	Ex-smoker
N21	+	Adenocarcinoma	IIA	70	M	Ex-smoker
N22	-	N/A	N/A	49	M	Never smoker
N25	+	Adenocarcinoma	IV	86	F	Never smoker
N27	+	Large Cell Ca.	IIB	73	M	Ex-smoker
N29	+	Adenocarcinoma	IA	84	F	Ex-smoker
N31	+	Adenocarcinoma	IA	72	F	Current smoker
N32	+	Large cell Ca.	IIA	62	M	Current smoker
N33	-	N/A	N/A	50	M	Never smoker
N36	+	NSCLC NOS	IIA	61	F	Ex-smoker
N37	+	Adenocarcinoma	IA	82	M	Ex-smoker
N40*	+	Adenocarcinoma	IIA	46	M	Never smoker
N41	+	Adenocarcinoma	IA	66	F	Current smoker
N45	-	N/A	N/A	49	M	Current smoker
N50*	+	Adenocarcinoma	IB	74	M	Ex-smoker
N51	-	N/A	N/A	60	M	Current smoker
N58	+	Adenocarcinoma	IIA	72	M	Ex-smoker
N64	-	N/A	N/A	30	M	Never smoker
N65	-	N/A	N/A	30	F	Ex-smoker
N74	+	Adenocarcinoma	IIB	59	F	Ex-smoker
Retrospective Patients						
N99	+	NSCLC NOS	IIB	59	M	Current smoker
N98	+	NSCLC NOS	IV	60	M	Ex-smoker
N97	+	Squamous Cell Ca.	IIIB	59	M	Ex-smoker
N96	+	Adenocarcinoma	IIA	83	M	Ex-smoker
N95	-	N/A	N/A	48	M	Never smoker
N94	-	N/A	N/A	49	M	Never smoker
N93	-	N/A	N/A	48	M	Current smoker
N92	-	N/A	N/A	29	M	Never smoker

Note: NSCLC cases with false negative results are annotated with an asterisk.

FIGURE 1. Duplicate amplification of two miRNAs (miRNA-21 and 210) and a small RNA (U6) for three separate sputum samples using qRT-PCR, demonstrating experimental reproducibility.

TABLE 2. Mean C_T and RN for the reproducibility experiments demonstrated in Figure 1

miRNA	C_T	Relative miRNA expression (Rn)
miR-21	20.75093 ± 0.08093	0.54175 ± 0.03033
miR210	35.98860 ± 0.00650	0.64605 ± 0.00289
U6	24.98653 ± 0.04730	1.00000

Data is mean $\pm$ standard deviation

FIGURE 2. Dendrogram displaying the average linkage clustering as performed with un-centered correlation metric between the 11 miRNA. The five miRNAs annotated with an asterisk were selected for prospective validation based on their ability to accurately categorize the retrospective cohort.

analysis method available in SPSS version 13 (IBM SPSS Inc., USA). The present study is a clinical validation study for our analytical methodology.

Results

Patient population

Table 1 displays the prospective and retrospective study participant case numbers with their respective NSCLC status and clinical features. Twenty-four patients were NSCLC-positive based on pathologically confirmed diagnosis, while six were negative controls.

MiRNA expression reproducibility

The experimental reproducibility miRNA levels play a critical role in evaluating the use of expression profiling as a potential screening tool. As such, two miRNAs (miR-21 and miR-210) and the endogenous control (U6) were selected with different CT (from lower limit 18 to higher limit 35) to test RT-qPCR technical reproducibility. Figure 1 demonstrates sample amplification curves for miR-21, miR-210 and U6 in three separate sputum samples performed in duplicate for each sample. Minimal variation was seen in each duplicate analysis highlighting the analytic reproducibility of the tests. Table 2 displays the mean Rn and CT values and associated standard deviations. The deviations of relative miRNA expressions ranged from 0.5% to 6.0 %, indicating that sputum miRNA expression with RT-qPCR is reliable and reproducible. RT-qPCR is ex-

tremely sensitive and subject to systematic errors. An observed standard error ranging from 0.25 to 0.5 for CT is considered to be an acceptable tolerance for RT-qPCR [35].

MiRNA panel optimization

Figure 2 displays the hierarchical cluster analysis dendrogram using the selected 11 miRNAs based on the eight retrospective patients (four NSCLC positive cases and four negative controls). Average linkage clustering was performed with an uncentred correlation metric between groups (i.e., the shorter the linkage distance the greater the degree of subject relatedness). Using the results of the cluster analysis for the 11 miRNAs on our eight retrospective patients (Figure 2), a different combination of miRNAs were tested using cluster analysis with a “trial-and-error” approach in order to get the best discrimination between cases and controls amongst the retrospective patients. Due to their high degree of relatedness, the miRNAs labelled with an asterisk (miR-21, miR-155, miR-210, miR-143, miR372) were then selected as a distinct miRNAs profile for further validation amongst the 30 prospectively accrued patients.

Table 3 displays the similarity matrix of cluster analysis for the eight retrospective patients’ miRNA profiles based on the relative expressions of the five selected miRNAs. The similarity matrix displays the correlation between vectors of values. The closer the value is to 1 in each cell, the greater the similarity between the two cases and the closer they will cluster in the dendrogram. Conversely, the closer the value is to 0, the more distinct the two cases will be from one another, a difference which is represented by a greater line length from the branch-

ing. Two clear clusters emerged accurately distinguishing the NSCLC positive cases from the NSCLC negative controls.

Optimized five miRNA (miR-21, miR-155, miR-210, miR-143, miR372) panel

A total of 30 prospectively accrued patients (twenty-four NSCLC positive cases, and six negative controls) had their sputum expression profiles analyzed in a double-blinded fashion. Hierarchical cluster analysis was performed using the previously selected five miRNA panel. Figure 3 displays the dendrogram for the 30 prospectively accrued patients. The NSCLC positive cases were clearly delineated from the NSCLC negative controls. The five miRNA expression profiles for the 30 patients have a high correlation between vectors of values, resulting in a sensitivity of 83.3% and specificity of 100% for the presence of NSCLC.

Discussion

MiRNA alterations are involved in the initiation, progression and metastases of human cancers [14, 17, 18]. The main molecular alterations are represented by variations in gene expression, usually mild, with consequences for a vast number of target protein coding genes. Several potential reasons for the widespread differential expression of miRNA in malignant compared with normal cells include: the location of these RNA species in cancer-associated genomic regions, epigenetic mechanisms, and alterations in the miRNA processing machinery [17]. MiRNA expression profiling of human tumors

Table 3. Similarity matrix for the eight retrospective cases based on five miRNA gene profiling with cluster analysis.

Case	Correlation between vectors of value							
	1:N99	2:N98	3:N95	4:N97	5:N96	6:N94	7:N93	8:N92
1:N99	1.000	.993	.322	.900	.926	-.266	.510	.072
2:N98	.993	1.000	.418	.893	.936	-.162	.601	.178
3:N95	.322	.418	1.000	.461	.530	.825	.972	.965
4:N97	.900	.893	.461	1.000	.947	-.083	.572	.216
5:N96	.926	.936	.530	.947	1.000	-.025	.658	.293
6:N94	-.266	-.162	.825	-.083	-.025	1.000	.691	.942
7:N93	.510	.601	.972	.572	.658	.691	1.000	.893
8:N92	.072	.178	.965	.216	.293	.942	.893	1.000

Five gene profile comprised of miR-21, miR-155, miR-210, miR-143, miR372.

Note: Values closer to 1 represent a higher degree of relatedness.

FIGURE 3. Dendrogram of the hierarchical cluster analysis of the prospectively accrued patients (24 NSCLC positive cases and six negative controls). Four false negative cases (N12, N14, N40 and N50) were observed. The overall sensitivity and specificity obtained using the five miRNA panel was 83.3% and 100%, respectively.

has identified signatures associated with diagnosis, staging, progression, prognosis and response to treatment [16-17, 20, 22-24, 26]. In addition, profiling has been exploited to identify miRNAs that may represent downstream targets of activated oncogenic pathways [17].

Although miRNAs seem to play an important role in carcinogenesis, the exact mechanism of influence of miRNAs over cancers is not fully understood. MiRNA profiling in cancers have generally been performed utilizing samples obtained from solid tumors. Since there may be potential differences between the miRNA profiles obtained from tumor tissue versus sputum samples from a cancer patient, a group of miRNAs that were known to be associated with malignancy or NSCLC in either blood, tissue, or both, including miR-21, miR-205, miRNA-210, let-7a and miR-17-5p, were selected for this study [14-15, 21]. The miRNAs in our selected panel were chosen due to their known association with NSCLC include: miR-155, miR-182, miR-21 [16, 20-22]. Since our aim is to develop a clinically useful and affordable screening test for NSCLC, we believe that an ideal miRNA panel should be comprised of three to five different miRNAs. At the time of inception of this study in 2009, there were no published studies that examined the feasibility of measuring miRNA from sputum as a means to detect NSCLC. Since that time, several groups, including Jiang et al. [9] have carried out studies regarding miRNA profiling using sputum. Jiang et al. selected a maximum of seven miRNAs and compared their exact expression levels of each miRNA as a means to discriminate lung cancer cases from normal controls [24-26]. This proved difficult since individual exact miRNA expression levels in sputum may be altered by a number of factors such as sample biovariation or RT-qPCR measurements that contain multiple steps. In contrast, in our study, miRNA profiling using analytic integrated model-recognition of unsupervised clusters supplanted the exact individual miRNA expression values.

MiRNAs have recently emerged as important regulators of gene expression. High-throughput analyses have shown that miRNA expression is commonly dysregulated in a variety of cancers. MiR-21 has been reported to be over-expressed in a number of cancers including breast, lung, colon, pancreas, and prostate. Hence, miR-21 expression was chosen for use in our panel to differentiate malignant from non-malignant tissue [8, 20, 21]. It is unreliable to use a single miRNA as a specific biomarker for lung cancer screening [14, 16, 23, 30]. To date, forty-three miRNAs related to lung cancer have been identified [15]. We used an initial 11 miRNA panel, based on the current literature evidence. MiR-17-5p has been shown to be over-expressed in breast, colon, lung, pancreas and prostate cancers

[16, 20-22]. MiR-143 and miR-145 are over-expressed in multiple cancer types including NSCLC [13, 15-16]. Let-7a-2 is down-regulated in lung cancer and breast cancer [13, 16]. MiR-155 is up-regulated in breast, colon and lung cancer whereas miR-92 is down-regulated in multiple solid tumors [13, 27]. MiR-205, a-let-7a, miR-372 and miR-182 over-expression have shown strong correlation with the presence of NSCLC, and may also portend patient prognosis [13, 21-23]. Multiple miRNAs enhance the sensitivity and specificity required of a potential screening tool; however, further evidence is required to determine the optimal panel for a given geographic region. Furthermore, the effect of benign disease processes on miRNA expression is not fully understood and may impact the predictive value of any miRNA panel.

Our study has demonstrated that our selected five miRNA panel (miR-21, miR-155, miR-210, miR-143, miR372) could provide a specific molecular signature for the screening of NSCLC based on analysis of sputum. Previous studies have demonstrated comparatively low sensitivities using sputum cytology (40%) or miR-21 expression alone (69%) as potential screening tools [7-9]. The 83.3% sensitivity and 100% specificity observed by our optimized five miRNA panel using hierarchical cluster analysis provides promise as a future screening test for NSCLC, for which no effective means of early detection currently exists. Comparatively, screening mammography demonstrated sensitivities around 74% [36]. Furthermore, the use of miRNA obtained from sputum as opposed to serum has the advantage of being less invasive to patients and more specific to pulmonary pathology. These results demonstrate the future promise that miRNA based molecular profiling may play a role in the early diagnosis of a variety of malignancies.

In this study, RT-qPCR technology was used to determine the miRNA expression levels in sputum samples. Based on our previous work using RT-qPCR, the miRNA expression can be detected from the genetic material from the equivalent of as few as three to 16 cells [27]. One concern with using sputum in widespread screening is that RNA may rapidly decay secondary to the presence of high levels of RNase activity; however, it has been previously reported that human sputum miRNAs may exist in a form that is resistant to RNase activity [9]. Furthermore, Patrick et al. demonstrated that miRNA can be detected in biological fluid even in the absence of any cancer cells [37]. This diminishes the importance of the "quality" of obtained sputum and its perceived effect on the efficacy of miRNA expression as a potential NSCLC screening tool.

Our data using the cluster analysis of an optimized five miRNA profile panel provides promise as a potential NSCLC screening test. Our clinical data validates our methodology of

cluster analysis as patients with NSCLC were clearly and reliably differentiated from NSCLC-negative controls.

The current study was conducted to assess the feasibility of quantifying sputum miRNAs, employing RT-qPCR and cluster analysis on the relative miRNA expressions to assess the methodological robustness of this potential screening method for the early detection of NSCLC. Given that various patient and tumor factors such as age, gender, smoking status, histological subtype of NSCLC, or comorbidities may alter the expression levels of an individual miRNA, we are currently testing this methodology in a prospective manner to a larger matched case control cohort of patients in the clinical setting whereby we will control for known confounding patient and disease factors and therefore be able establish confidence interval ranges for both NSCLC positive and negative patients. The ultimate long-term goal of our miRNA profiling research using cluster analysis is to develop a robust, affordable, accurate and non-invasive screening test for the early detection of NSCLC for use in patients at high risk of developing NSCLC such as current or previous smokers.

Sources of Support

Alberta Health and Wellness (equipment support).

References

- Jemal A, Siegel R, Ward E, Ward E, Hao Y, Xu J, Murray T, Thun MJ. Cancer Statistics, 2008. *CA Cancer J Clin* 2008; 58(2):71-96.
- Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. *Ann Oncol* 2007; 18(3):581-592.
- Ettinger DS, Akerley W, Bepler G, Blum MG, Chang A, Cheney RT, et al. Non-small cell lung cancer clinical practice guidelines in oncology. *J Natl Compr Canc Netw* 2010; 8(7):740-801
- Yang W, Kaur D, Okayama Y, Ito A, Wardlaw AJ, Brightling CE, Bradding P. Human lung mast cells adhere to human airway smooth muscle, in part, via tumor suppressor in lung cancer-1. *J Immunol* 2006; 176(2):1238-1243.
- Lacroix J, Becker HD, Woerner SM, Rittgen W, Drings P, von Knebel Doeberitz M. Sensitive Detection of Rare Cancer Cells in Sputum and Peripheral Blood Samples of Patients with Lung Cancer by Prepro-GRP Specific RT-PCR. *Int J Cancer* 2001; 92(1):1-8.
- Li R, Todd NW, Qiu Q, Fan T, Zhao RY, 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(2 Pt 1):482-487.
- Wu GP, Wang EH, Li JH, Fu ZM, Hans S. Clinical application of the liquid-based cytological test in cytological screening of sputum for the diagnosis of lung cancer. *Respirology* 2009; 14(1):124-128.
- Byers T, Wolf HJ, Franklin WA, Braudrick S, Merrick DT, et al.. Sputum cytologic atypia predicts incident lung cancer: defining latency and histologic specificity. *Cancer Epidemiol Biomarkers Prev* 2008; 17(1):158-162
- Xie Y, Todd NW, Liu Z, Zhan M, Fang H, et al. Altered miRNA expression in sputum for diagnosis of non-small cell lung cancer. *Lung Cancer* 2010; 67(2):170-176.
- Keohavong P, Gao WM, Zheng KC, Mady H, Lan Q, Melhem M, Mumford J. Detection of K-ras and p53 mutations in sputum samples of lung cancer patients using laser capture microdissection microscope and mutation analysis. *Anal Biochem* 2004; 324(1):92-99.
- Chanin TD, Merrick DT, Franklin WA, Hirsch FR. Recent developments in biomarkers for the early detection of lung cancer: perspectives based on publications 2003 to present. *Curr Opin Pulm Med* 2004; 10(4):242-247.
- Deng S, Calin GA, Croce CM, Coukos G, Zhang L. Mechanisms of microRNA deregulation in human cancer. *Cell Cycle* 2008; 7(17):2643-2646.
- Mallick R, Patnaik SK, Yendamuri S. MicroRNAs and lung cancer: biology and applications in diagnosis and prognosis. *J Carcinog* 2010; 9:8.
- Lu J, Getz G, Miska EA, Alvarez-Saavedra E, Lamb J, et al. MicroRNA expression profiles classify human cancers. *Nature* 2005; 435(7043):834-838.
- Rosenfeld N, Aharonov R, Meiri E, Rosenwald S, Spector Y, et al. MicroRNAs accurately identify cancer tissue origin. *Nat Biotechnol* 2008; 26(4):462-469.
- Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* 2006; 9(3):189-198.
- Calin GA, Croce CM. MicroRNA-cancer connection: the beginning of a new tale. *Cancer Res* 2006; 66(15):7390-7394.
- Lynam-Lennon N, Maher SG, Reynolds JV. The roles of microRNA in cancer and apoptosis. *Biol Rev Camb Philos Soc* 2009; 84(1):55-71.
- Chen X, Ba Y, Ma L, Cai X, Yin Y, et al. Characterization of microRNAs in serum: a novel class of biomarkers for diagnosis of cancer and other diseases. *Cell Res* 2008; 18(10):997-1000.
- Rabinowits G, Gercel-Taylor C, Day JM, Taylor DD, Kloecker GH. Exosomal microRNA: a diagnostic marker for lung cancer. *Clin Lung Cancer* 2009; 10(1):42-46.
- Si ML, Zhu S, Wu H, Lu Z, Wu F, Mo YY. miR-21-mediated tumor growth. *Oncogene* 2007; 26(19):2799-2803.
- Markou A, Tsaroucha EG, Kaklamanis L, Fotinou M, Georgoulas V, Lianidou ES. Prognostic value of mature microRNA-21 and microRNA-205 over-expression in non-small cell lung cancer by quantitative real-time RT-PCR. *Clin Chem* 2008; 54(10):1696-1704.

23. Jay C, Nemunaitis J, Chen P, Fulgham P, Tong AW. MiRNA profiling for diagnosis and prognosis of human cancer. *DNA & Cell Biology* 2007; 26(5):293-300.
24. Yu L, Todd NW, Xing L, Xie Y, Zhang H, et al. Early detection of lung adenocarcinoma in sputum by a panel of microRNA markers. *Int J Cancer* 2010; 127 (12): 2870-2878.
25. Xing L, Todd NW, Yu L, Fang H, Jiang F. Early detection of squamous cell lung cancer in sputum by a panel of microRNA markers. *Mod Pathol*. 2010 Aug; (23(8): 1157-64.
26. Yu SL, Chen HY, Chang GC, Chen CY, Chen HW, et al. MicroRNA signature predicts survival and relapse in lung cancer. *Cancer Cell* 2008; 13(1):48-57.
27. Wilson R, Brunet B, Guo L, Amanie J, Fairchild A, Gabos Z, Nijjar T, Scrimger R, Yee D, Xing J. Identification of a new microRNA expression profile as a potential cancer screening tool. *Clin Invest Med* 2010; 33(2):124-132.
28. Schmittgen TD, Lee EJ, Jiang J, Sarkar A, Yang L, Elton TS, Chen C. Real-time PCR quantification of precursor and mature microRNA. *Methods* 2008; 44(1):31-38.
29. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* 2001; 25(4):402-408.
30. Riffenburgh RH. *Statistics in medicine*. Second Edition, Academic Press 2006; 524.
31. Hastie T, Tibshirani R, Friedman J. *The elements of statistical learning*, Second Edition. New York. Springer 2009; 2:520-528.
32. Peltier HJ, Latham GJ. Normalization of microRNA expression levels in quantitative RT-PCR assays: identification of suitable reference RNA targets in normal and cancerous human solid tissues. *RNA* 2008; 14(5):844-852.
33. Michael MZ, O'Connor SM, van Holst Pellekaan NG, Young GP, James RJ. Reduced accumulation of specific microRNAs in colorectal neoplasia. *Mol Cancer Res* 2003; 1(12):882-891.
34. Lorio MV, Ferracin M, Liu CG, Veronese A, Spizzo R, et al. MicroRNA gene expression deregulation in human breast cancer. *Cancer Res* 2005; 65(16):7065-7070.
35. Becker C, Hammerle-Fickinger A, Riedmaier I, Pfaffl MW. mRNA and microRNA quality control for Rt-qPCR analysis. *Methods* 2010; 50(4):237-243.
36. Gøtzsche PC, Nielsen M. Screening for breast cancer with mammography. *Cochrane Database Syst Rev* 2011; 1:CD001877.
37. Mitchell PS, Parkin RK, Kroh EM, Fritz BR, Wyman SK, et al. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci USA* 2008; 105(30):10513-10518.
38. Witt BL, Wallander ML, Layfield LJ, Hirschowitz S. Respiratory cytology in the era of molecular diagnostics: a review. *Diagn Cytopathol*. 2012 Jun; 40(6): 556-563.
39. Heighway J, Betticher DC. Lung: Non-small cell carcinoma. *Atlas GenetCytogenet Oncol Haematol*. February 2004. Available at: <http://AtlasGeneticsOncology.org/Tumors/LungNonSmallCellID5141.html>. Accessed August 14, 2008.