Unique microRNA-profiles in *EGFR*-mutated lung adenocarcinomas

Maria Moksnes Bjaanæs[1,2], Ann Rita Halvorsen[1], Steinar Solberg[3], Lars Jørgensen[3], Tommaso A. Dragani[4], Antonella Galvan[4], Francesca Colombo[4], Marco Anderlini[4], Ugo Pastorino[5], Elin Kure[1], Anne-Lise Børresen-Dale[1,6], Odd Terje Brustugun[1,2,6] and Åslaug Helland[1,2,6]

[1] Department of Genetics, Institute for Cancer Research, Oslo University Hospital-The Norwegian Radium Hospital, Oslo, Norway

[2] Department of Oncology, Oslo University Hospital-The Norwegian Radium Hospital, Oslo, Norway

[3] Department of Cardiothoracic Surgery, Oslo University Hospital-Rikshospitalet, Oslo, Norway

[4] Department of Predictive and Preventive Medicine, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[5] Department of Thoracic Surgery, Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy

[6] Institute of Clinical Medicine, Faculty of Medicine, University of Oslo, Oslo, Norway

The findings of mutations and the development of targeted therapies have improved lung cancer management. Still, the prognosis remains poor, and we need to know more about the genetic and epigenetic alterations in lung cancer. MicroRNAs are involved in crucial biological processes like carcinogenesis by regulating gene expression at the post-transcriptional level. In this project, we have studied the microRNA expression of lung adenocarcinomas and corresponding normal lung tissue and correlated the expression with clinical data and EGFR- and KRAS-mutational status. Agilent microarrays have been used, examining microRNA expression in 154 surgically resected lung adenocarcinomas and 20 corresponding normal lung tissue samples. Findings were confirmed by RT-qPCR in the same cohort and in an independent cohort of 103 lung cancer patients. EGFR and KRAS mutation analyses were also performed. 129 microRNAs were significantly differentially expressed in lung adenocarcinomas compared with normal lung tissue, and 17 microRNAs were differentially expressed between EGFR-mutated and EGFR wildtype tumors. We identified microRNAs associated with time to progression. We have identified several aberrantly expressed microRNAs that discriminate lung adenocarcinomas from normal lung tissue, and hence may be potential biomarkers for early detection. We have found microRNAs that are differentially expressed between EGFR-mutated and EGFR wildtype lung adenocarcinomas, suggesting that microRNAs can be used as molecular biomarkers in classification. We hypothesize that microRNA expression can be used as biomarkers for clinical course.

Lung cancer is the most common cause of cancer deaths worldwide.[1] Non-small cell lung cancer (NSCLC) accounts for approximately 85% of all diagnosed lung cancers, and varies both in molecular and clinical presentation. Adenocarcinomas and squamous cell carcinomas are main histological subtypes of NSCLC, and adenocarcinomas are increasing in incidence both in Norway[2] and worldwide.[3]

The identification of key oncogenic drivers has led to a more personalized approach in the treatment of advanced disease and has indicated a paradigm shift in the

What's new?
Precise molecular mechanisms for the altered expression of microRNAs in lung cancer are unclear. Here, the authors analyzed microRNA expression in a large sample set of lung adenocarcinomas and corresponding normal lung tissue. They identified several aberrantly expressed microRNAs that discriminate lung adenocarcinomas from normal tissue and are thus potential biomarkers for early detection. They found microRNAs differentially expressed between EGFR-mutated and EGFR wild-type lung adenocarcinomas, suggesting that microRNAs can be used as molecular biomarkers in classification. Moreover, miR-500a* expression was associated with time to progression. This study suggests that microRNA expression can be used as biomarkers for clinical course.

understanding of the NSCLC biology. The discovery of activating mutations in the *EGFR* gene implicating sensitivity to EGFR tyrosine kinase inhibitors has been important for the care of NSCLC patients.

The *KRAS* mutation is considered to be one of the most frequent mutations in lung adenocarcinomas[4] and is more common in adenocarcinomas arising in smoking patients. Drugs targeting the KRAS pathway are currently being tested in clinical studies.

Despite the development of targeted therapy and advances in surgery, chemotherapy and radiotherapy over the last decades, the death rate from lung cancer has remained largely unchanged. Because of the overall poor prognosis, new understanding of the lung cancer biology and new treatment strategies are urgently needed.[5]

MicroRNAs are thought to play an essential role in the development and progression of human malignancies. MicroRNAs are small noncoding RNAs that regulate gene expression by mRNA degradation and translational suppression.[6] Each microRNA is predicted to target hundreds of mRNAs, and 10–30% of all protein-encoding human genes may be regulated by these mechanisms.[7,8] MicroRNAs are expressed in a tissue-specific manner and are deregulated in a variety of cancers, playing a key role in tumorigenesis.[9–11]

MicroRNA expression can distinguish tumors from normal tissue and tumors with different developmental origins.[12] NSCLC tissue has differentially expressed microRNAs compared with normal lung tissue,[11,13] and some of the published microRNA expression studies in NSCLC are reviewed by Guan *et al.*[14] Little is known, however, about differences in microRNA expression in lung adenocarcinomas with different *EGFR*- and *KRAS*-mutational status.

We conducted an explorative microRNA expression study in 154 surgically resected lung adenocarcinomas and 20 corresponding normal lung tissue samples, using microRNA microarrays. Our main focus was to investigate how the microRNA expression is different in lung adenocarcinoma tissue from normal lung tissue and to find microRNAs associated with the clinical important *EGFR* mutation status. We have also investigated how microRNA expression co-varies with progression-free survival. Expression of selected microRNAs was confirmed using RT-qPCR in the same cohort and then validated in an independent sample set of 103 paired lung adenocarcinomas and normal lung tissue samples.

Material and Methods
Patients and tissue samples
The participants in our study were patients with operable lung cancer who were admitted to the cardiothoracic surgery department at Oslo University Hospital–Rikshospitalet from 2006 to 2011. They received oral and written information and signed a written consent form before entering the project.

The histopathological and clinical data were collected from the hospital and follow-up information from questionnaires and the patients' local hospitals.

Only one patient had received radio- and chemotherapy prior to surgery. Forty-two patients received adjuvant chemotherapy following standard guidelines. Some patients did not receive all four cycles of chemotherapy due to side-effects, and these are not included in this number. Patients over 70 years were not offered chemotherapy.

Tumor tissue and normal lung tissue were dissected and prepared immediately after the surgical specimen had been removed from the patient. The tumor tissue was dissected from the tumor's periphery, with presumably vital tissue without necrosis. The normal lung tissue was collected from the resected lung or lobe, at least 10 cm from the macroscopic tumor. Tissue specimens were snap frozen in liquid nitrogen and stored at −80°C until RNA isolation.

The project was approved by the institutional review board and Regional Ethics Committee (S-05307). The clinical and pathological characteristics are shown in Table 1.

For validation of selected microRNAs, an independent sample set of 103 lung adenocarcinomas and paired normal lung tissue samples were used, available at Fondazione IRCCS Istituto Nazionale dei Tumori, Milan, Italy. All patients underwent lung lobectomy. At the end of surgery, a section from the tumor and a small section of non-involved lung tissue were harvested and stored frozen or placed directly in RNAlater solution (Life Technologies, Grand Island). Information on histological diagnoses (made by the Pathology Departments of the recruiting institute or hospital) was retrieved from the clinical records and data regarding sex, age at diagnosis and clinical stage was recorded when the samples were taken. Patients' characteristics for the validation population are detailed in Supporting Information Table S1. Each subject gave informed consent to the use of biological samples for research purposes.

Cancer Genetics

Table 1. Characteristics of patients and tumors included in the microarray analyses

Variable	Tumor samples	Normal lung samples
	n = 154	*n* = 20
Age (years)		
Mean	66.2	66
Median	66.5	68.1
<65	69	9
>65	85	11
Sex		
Females	87	14
Males	67	6
Smoking history		
Current	87	7
Former	48	7
Never	19	6
EGFR mutation status		
EGFR mutated	22	
EGFR wt	130	
EGFR not tested	2	
KRAS mutation status		
KRAS mutated	52	
KRAS wt	96	
KRAS not tested	6	
Stage		
Ia	45	
Ib	46	
IIa	24	
IIb	12	
IIIa	26	
IV	1	
Adjuvant chemotherapy		
Yes	42	
No	109	
Not known	3	
Status		
1: Alive with no disease	103	
2: Alive with disease	18	
3: Dead with disease	0	
4: Dead with no disease	9	
5: Dead because of disease	24	
Follow up time (months)		
Mean	31.7	
Range	7−63	

RNA extraction

Standard TRIZOL methods (Invitrogen, Carlsbad, CA), as specified by the manufacturer's instructions, were used to extract total RNA from the snap-frozen tumor and normal lung tissue. RNA quantity and quality (yield, 260/280 ratio and 260/230 ratio) were determined using the NanoDrop ND-1000 spectrometer (NanoDrop Technologies), and RNA integrity numbers (RIN) were measured with the use of the 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA) according to the manufacturer's protocol.

MicroRNA expression

MicroRNA microarrays from Agilent Technologies (Agilent human microRNA microarray kit release 16.0, 8 × 60 K) were used for microRNA profiling. This microarray encodes 1,205 human microRNAs and 144 human viral microRNAs listed in the Sanger miRBase (Release 16.0).

300 ng of total RNA was labeled with Cy3 with ligation of a RNA-linker. The slides were incubated with the labeled RNA for 20 hr at 55°C, and then the slides were washed. Arrays were scanned using Agilent Microarray Scanner (Agilent Technologies, Santa Clara, CA). The raw data were pre-processed with Agilent's Feature Extraction Software, where default parameters were employed (Agilent feature extraction v. 10.7.3.1).

The microRNA profiling was done on 200 samples. Samples with an error in the QC reports generated by the feature extraction procedure were run again, and samples with an error in the QC report after repeated runs (5 samples) were excluded from further analysis. Twenty-one samples were excluded from further investigations due to expression of very few microRNAs. A total of 174 samples (154 lung adenocarcinomas and 20 normal lung tissue samples) remained in the analyses.

We filtered out microRNAs that were detected in less than 10% of the samples, with remaining 570 microRNAs for further analysis.

EGFR mutation analyses

Mutation analyses of *EGFR* exons 18–21 were performed on 152 of the tumor samples using the TheraScreen EGFR mutation kit (DxS, Manchester, UK). The assay is designed to detect 28 specific mutations in the *EGFR* gene by real-time PCR. Assays were carried out according to the manufacturer's protocol and with the use of the Roche LightCycler 480 real-time PCR system. Data analyses were performed by employing the LightCycler Adapt software (LightCycler 480 Software, v. 1.5). Some of the results were previously published by Helland *et al.*[15]

KRAS mutation analysis

We used the wobble-enhanced ARMS (WE-ARMS) method[16] for detecting *KRAS* mutations in 148 lung adenocarcinoma samples. This mutation assay detects the seven most

commonly reported mutations in the *KRAS* gene—KRAS g.34G>C (p.G12R), g.34G>A (p.G12S), g.34G>T (p.G12C), g.35G>A (p.G12D), g.35G>C (p.G12A), g.35G>T (p.G12V) and g.38G>A (p.G13D)—by real-time PCR.

Reverse transcriptase quantitative polymerase chain reaction (RT-qPCR) for microRNA expression

Microarray expression of selected microRNAs was validated by RT-qPCR, both in the same cohort and in the validation cohort as described. The RT-qPCR was done using a microRNA-specific TaqMan microRNA assay kit (Applied Biosystems, Foster City, CA) and Applied Biosystems 7900H Fast Real-Time system. Reverse transcription of 10 ng of RNA was performed by employing the Taqman microRNA reverse transcription kit and microRNA-specific primers (Applied Biosystems) (Supporting Information Table S2).

The cDNA from the reverse transcription reaction was added to the TaqMan Universal Master Mix II no UNG and the labeled microRNA specific TaqMan probes (Applied Biosystems). RNU 6B, small nuclear RNAs, were used as endogenous controls for normalization of input cDNA levels (Applied Biosystems). We used a commercial Ambion breast control (Life Technologies, http://www.lifetechnologies.com/) as a calibrator.

The RT-qPCR was performed on all samples in duplicate. MicroRNA expression was quantified as $\Delta\Delta C_t$ values, where C_t = threshold cycle, $\Delta C_t = (C_t$ target microRNA $- C_t$ RNU 6B) and $\Delta\Delta C_t = (\Delta C_t$ target microRNA $- \Delta C_t$ calibrator), and fold change for each microRNA was calculated by the $2^{-\Delta\Delta C_t}$ method.[17] This was done using RQ manager software, v. 1.2 (Applied Biosystems).

Statistical analyses

Normalization and explorative analysis. The normalization and quality control of microarray data was carried out using the Genespring GX Analysis Software v.12.1 (Agilent Technology). Data were preprocessed by $\log_2$ transformation, and the normalization between all arrays was done by the 90th percentile method.

J express software[18] was used to perform unsupervised hierarchical clustering on both samples and microRNAs. This was done to investigate different associations between technical and clinicopathological variables. Pearson correlation (distance measure) and average linkage (WPGMA) were used in the hierarchical clustering.

Differentially expressed microRNAs. Examining differentially expressed microRNAs between groups of samples was done using significant analysis of microarray (SAM)[19] in the J express software.[18] The SAM analysis was applied to assess the difference in the expression levels of microRNAs between lung adenocarcinoma tissue and normal lung tissue, different clinical features and mutational status in the lung adenocarcinoma samples. The microRNAs with FDR <1% and fold change below −1.5 or above 1.5 were considered as significantly differentially expressed.

Survival analysis. Time to progression was calculated from the date of diagnosis to the date of diagnosis of local relapse, distant metastasis or lung cancer death. Patients that died of other causes or moved out of the country were treated as censored. One patient had distant metastasis (stage IV) at the time of surgery and was excluded from the survival analysis. Based on the follow-up-time in this study, we considered time to progression the most informative end-point.

We used a univariate Cox regression analysis to identify microRNA expression and clinicopathological factors (sex, stage, age at surgery, adjuvant chemotherapy, smoking status, tumor size and *EGFR-* and *KRAS*-mutational status) with significant influence on progression-free survival.

The Cox regression model was used for multivariate analyses. Factors included in the multivariate model were sex, stage, age at surgery, adjuvant chemotherapy, smoking status, tumor size and the expression of the microRNAs most related to time to progression in the univariate analysis. A *p* value <0.05 was considered as statistically significant. The survival analyses were done using SPSS (v. 18).

The RT-qPCR results were analyzed using the non-parametric Mann–Whitney test to evaluate statistically significant differences in microRNA expression between lung adenocarcinoma tissue compared with the normal lung tissue and the *EGFR*-mutated tumors compared with *EGFR* wild-type tumors. The Cox regression analysis was used for validating the survival data. The microRNAs expression values of the microRNAs associated with time to progression were dichotomized to high or low expression (based on the median expression value) and were tested with the log-rank test with Kaplan–Meier curves.

The RT-qPCR analysis in the independent validation cohort was done using Wilcoxon signed-rank tests (paired samples). *P* values <0.05 were considered statistically significant. The RT-qPCR results were analyzed using SPSS (v. 18).

Pathway analysis. We used the microRNA target filter in IPA (Ingenuity Pathway Analysis, Ingenuity® Systems, http://www.ingenuity.com) to search for genes associated with the differentially expressed microRNAs in our experiment. The microRNA target filter uses MiRecord and TarBase databases to search for validated microRNA target genes and Targetscan for target site prediction. We used the confidence filter and included the target genes that were experimentally observed or highly predicted to be associated with the specific microRNAs (March 2013).

Results

Patients and tumor characteristics

Clinicopathological characteristics are shown in Table 1. We identified 22 *EGFR* mutations in the 152 adenocarcinomas tested (14.5%). Seventeen of the 22 patients with *EGFR*-mutated tumors were women. Twelve of the patients with tumors harboring *EGFR* mutation were never smokers.

We identified 52 *KRAS*-mutated tumors among the 148 tumors tested (35.1%), affecting 18 males and 34 females. Only
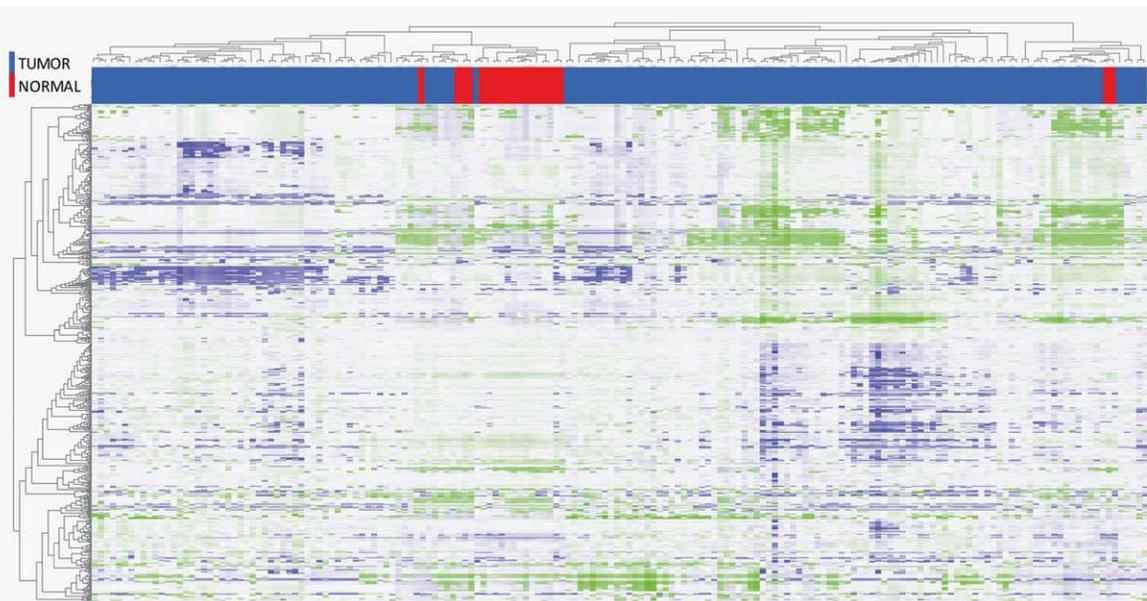
<div style="writing-mode: vertical">Cancer Genetics</div>

Figure 1. Unsupervised clustering of 570 microRNAs (rows) in 154 samples of lung adenocarcinoma tissue (colored blue in the column heading) and 20 samples of normal lung tissue (colored red in the column heading). The bricks that are colored green in the heatmap have a relatively high microRNA expression and the bricks that are colored blue have a relatively low microRNA expression. Distance measure: Pearson correlation. Linkage: Average (WPGMA).

three of the patients with *KRAS*-mutated tumors were never smokers. No tumors had both *EGFR*- and *KRAS*-mutation.

Analysis of tumor tissue compared with normal lung tissue

Unsupervised hierarchical clustering of all the 174 samples (tumor tissue and normal lung tissue) showed that most of the normal lung tissue clustered separately from the tumor tissue (Fig. 1). Eighteen of the 20 normal lung samples clustered together. The remaining 2 normal lung samples clustered together in a different cluster. The normal lung tissue samples did not cluster together with their corresponding tumor sample.

To identify microRNAs that were differentially expressed between the lung adenocarcinoma tissue and the normal lung tissue, we performed SAM analyses. We identified 129 microRNAs that were differentially expressed between tumor tissue compared with the normal lung tissue, with a fold change more than 1.5 or less than −1.5 and FDR <1% (Supporting Information Table S3). In the tumor tissue 95 microRNAs were downregulated, and 4 microRNAs (miR-133a, miR-4328, miR-187*, miR-204) were downregulated with more than a 20-fold difference. Thirty-four microRNAs were upregulated in tumor and 1 microRNA (miR-9*) was upregulated by more than 20-fold compared with the normal lung tissue.

When we compared the 20 paired tumor samples to their correspondent normal lung sample, we identified 129 microRNAs that were significantly differentially expressed. Ninety-five of the microRNAs overlapped with the 129 microRNAs that were differentially expressed when looking at all tumor samples compared with normal lung tissue samples. The

microRNAs with the highest and lowest fold change in the total group were still significantly differentially expressed in the paired samples.

Analysis of the tumor samples according to mutation status and clinical variables

Unsupervised hierarchical clustering of the tumor samples did not separate the clinical variables or mutation status into different clusters (Supporting Information Fig. S1).

Seventeen microRNAs were differentially expressed between *EGFR*-mutated tumors ($n = 22$) and *EGFR* wt tumors ($n = 130$), as determined by SAM analysis. Sixteen microRNAs were upregulated and one was downregulated in the *EGFR*-mutated tumors.

Only 3 microRNAs were differentially expressed between the 52 *KRAS*-mutated and 96 *KRAS* wt tumors. Two microRNAs were downregulated and one was significantly upregulated in the *KRAS*-mutated tumors. The microRNAs differentially expressed between *EGFR*-mutated and wt tumors and *KRAS*-mutated and wt tumors are shown in Table 2.

No major differences in the microRNA expression were identified when comparing the microRNA expression between different clinicopathological features like sex, age at the time of surgery, pathological stage and smoking history. Only one microRNA (miR-516-5p) was significantly differentially expressed between tumors from male and female participants. MiR-3137 was upregulated in the older patients (>65 years). Only a few microRNAs were differentially expressed between tumors from patients with different smoking status (never, former and current smokers).

Cancer Genetics

Pathway analysis

Using the microRNA target filter in IPA, we searched for genes involved in the signaling pathways that are experimentally observed or highly predicted to be regulated by the selected microRNAs.

Looking at the mRNA targets of the 129 differentially expressed microRNAs between lung adenocarcinomas and normal lung tissue, we identified 10,197 mRNAs targeted by 87 of the 129 differentially expressed microRNAs.

The 17 microRNAs that were differentially expressed between *EGFR*-mutated and *EGFR* wt tumors (from Table 2) were associated with 1,254 mRNA targets. Six microRNAs were associated with the EGFR signaling pathway (Table 3). Two microRNAs (miR-184 and miR-30d*) were experimentally confirmed to target AKT2 and CDK6, respectively.[20,21]

These genes are shown in Table 3 and the pathway in Figure 2.

Survival analysis

We used univariate Cox regression analysis to study clinical features related to progression-free survival. Tumor stage (I vs. II vs. III) was significantly associated with progression-free survival ($p = 0.019$). Tumor size was borderline significant ($p = 0.056$), and patients with early stage tumors tended to have a longer time to progression. Sex, age at diagnosis, adjuvant chemotherapy, smoking history and *EGFR*- or *KRAS* mutations were not significantly associated with progression-free survival in our dataset (Supporting Information Table S4).

When using univariate Cox regression analysis on the continuous expression levels to the 570 microRNAs, we identified 55 microRNAs that were associated with progression-free survival, with a p value <0.05. Expression-levels higher than median of miR-10a* and lower levels than median of miR-500a* were significantly related to shorter time to progression, with a p value ≤ 0.001. In the multivariate Cox proportional model, miR-10a* and miR-500a* were still significant (p value $= 0.002$ and 0.014, respectively) after adjusting for known clinical predictive factors (Supporting Information Table S5).

Validation of the microRNA results using quantitative reverse transcriptase polymerase chain reaction (RT-qPCR)

We selected six microRNAs (miR-133a, miR-9*, miR-148a*, miR-452, miR-10a* and miR-500a*) for technical validation by RT-qPCR, performed in the same cohort used for the microarray analyses. In addition, four microRNAs were validated in an independent validation cohort (miR-133a, miR-

Table 2. MicroRNAs that were significantly differentially expressed between *EGFR*-mutated compared with wt tumors and between *KRAS*-mutated compared with wt tumors

	EGFR-mutated vs. wt lung adenocarcinoma tumors		
	MicroRNA	**FDR**	**FC[1]**
1	hsa-miR-184	<0.01	5.538
2	hsa-miR-339-3p	<0.01	4.121
3	hsa-miR-148a*	<0.01	3.509
4	hsa-miR-224*	<0.01	5.09
5	hsa-miR-452	<0.01	4.554
6	hsa-miR-450a	<0.01	5.586
7	hsa-miR-423-3p	<0.01	3.873
8	hsa-miR-654-5p	<0.01	3.768
9	hsa-miR-532-5p	<0.01	1.613
10	hsa-miR-3607-5p	<0.01	4.969
11	hsa-miR-28-3p	<0.01	3.261
12	hsa-miR-30d*	<0.01	3.515
13	hsa-miR-532-3p	<0.01	1.949
14	hsa-miR-500a*	<0.01	1.706
15	hsa-miR-502-3p	<0.01	1.769
16	hsa-miR-605	<0.01	4.133
17	hsa-miR-492	<0.01	−3.466
	KRAS-mutated vs. wt lung adenocarcinoma tumors		
1	hsa-miR-371-5p	<0.01	−2.127
2	hsa-miR-564	<0.01	−2.526
3	hsa-miR-100	<0.01	1.573

Abbreviations: FDR, false discovery rate; FC, fold change.
[1]*EGFR*-mutated/*EGFR* wt or *KRAS*-mutated/*KRAS* wt.

Table 3. MicroRNAs differentially expressed between *EGFR*-mutated and wt adenocarcinomas of the lung and the predicted or observed target gene from IPA

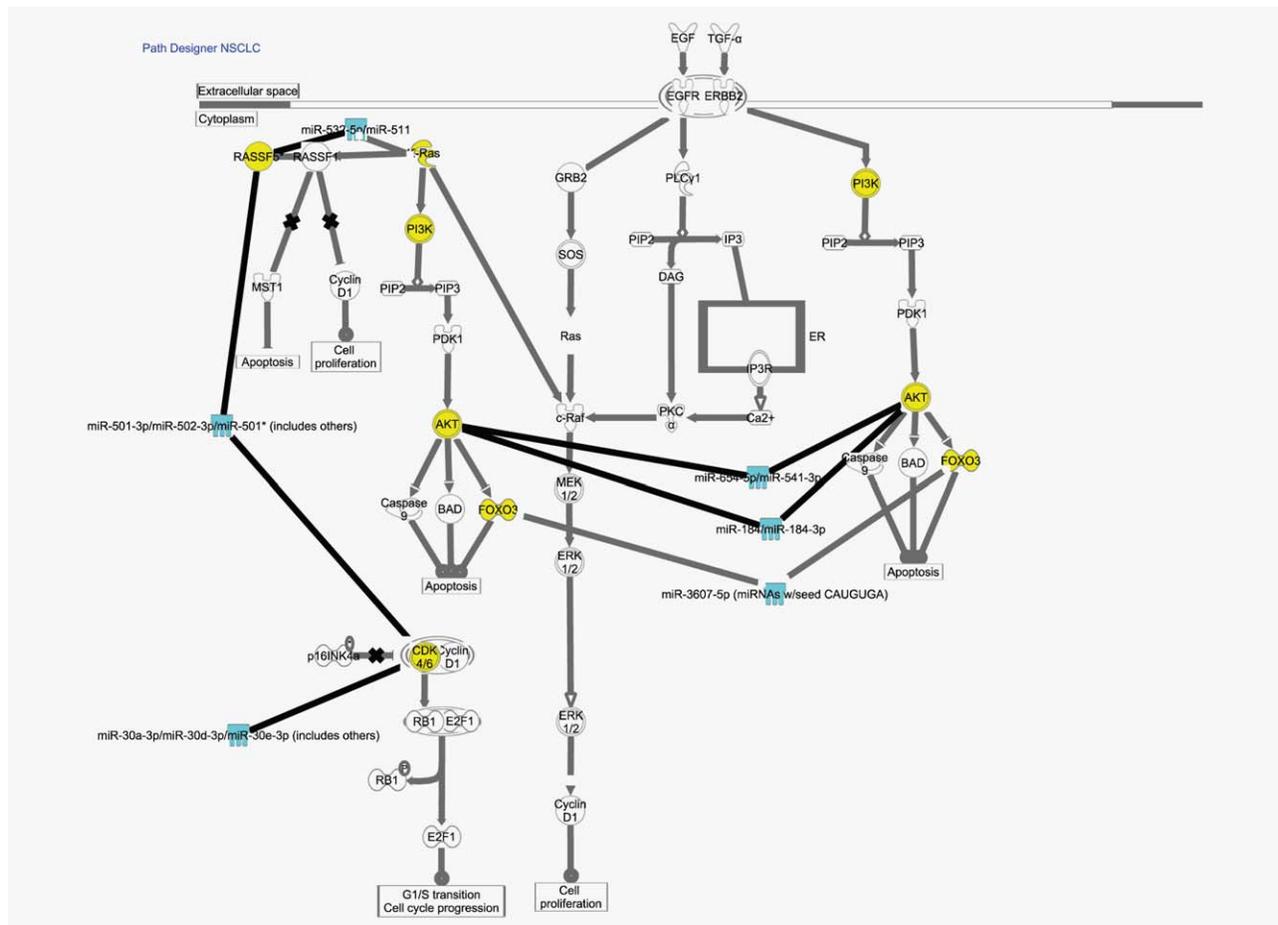| ID	Fold change	Source	Confidence	Gene symbol
MiR-184	5,538	miRecords	Experimentally observed	AKT2
MiR-30d*	3,515	TarBase, miRecords	Experimentally observed	CDK6
MiR-3607-5p	4,969	TargetScan human	High (predicted)	FOXO3
MiR-502-3p	1,769	TargetScan human	High (predicted)	CDK6
MiR-502-3p	1,769	TargetScan human	High (predicted)	RASSF5
MiR-532-5p	1,613	TargetScan human	High (predicted)	KRAS
MiR-532-5p	1,613	TargetScan human	High (predicted)	RASSF5
MiR-654-5p	3,768	TargetScan human	High (predicted)	AKT3
MiR-654-5p	3,768	TargetScan human	High (predicted)	PIK3R6

Cancer Genetics

Figure 2. EGFR pathway in NSCLC from the IPA analysis. The turquoise molecules are microRNAs that were differentially expressed between *EGFR*-mutated tumors and *EGFR* wt tumors in our experiment. The yellow molecules are genes that are experimentally observed or highly predicted to be regulated by the highlighted microRNAs.

9*, miR-21 and miR-126). MiR-133a, miR-9*, miR-21 and miR-126 were chosen because they showed significant differences (low FDR and high positive/negative fold change) between lung adenocarcinoma tissue compared with the normal lung tissue. MiR-9* and miR-133a are not previously known to be altered in lung adenocarcinomas and were the microRNAs with the highest and lowest fold changes (miR-9* FC = 29.1 and miR-133a FC = −28.8) in our study. We also wanted to confirm findings from other studies, and miR-21 and miR-126 were selected because they are known to be altered in lung cancer tissue. MiR-148a* and miR-452 were chosen because they showed significant differences between *EGFR*-mutant and wt adenocarcinoma tissue with a relatively high fold change, they were expressed in a large number of samples and they are located on different chromosomes (chr 7 and chr X). MiR-10a* and miR-500a* were the two microRNAs with the strongest association with time to progression, with a p value ≤ 0.001.

In the technical validation analyses in the same cohort, both miR-133a and miR-9* were differentially expressed between tumor tissue and normal lung tissue ($p < 0.001$). MiR-452 was significantly differentially expressed ($p = 0.009$)

between the *EGFR*-mutated and *EGFR* wt tumors analyzed when using RT-qPCR, but miR-148a* was not validated ($p = 0.264$). MiR-500a* was significantly associated with time to progression analyzed using a Cox regression analysis ($p = 0.03$). The low values of miR-500a* was significantly associated with poor outcome when the microRNA expression values were dichotomized to high or low expression and tested with the log-rank test with Kaplan–Meier curves ($p = 0.029$). MiR-10a* did not reach significance as a prognostic marker in the validation analysis ($p = 0.401$) (Fig. 3).

The four selected microRNAs that were differentially expressed between lung adenocarcinomas and matched normal lung tissue were analyzed in an independent validation cohort. All four microRNAs (miR-133a, miR-9*, miR-21 and miR-126) were significantly differentially expressed, with a p value <0.05 when using RT-qPCR, thus confirming the microarray results (Supporting Information Table S6).

Discussion

In this explorative study, we analyzed the microRNA expression in tumor tissue and normal lung tissue from patients with primary operable lung adenocarcinomas. We identified
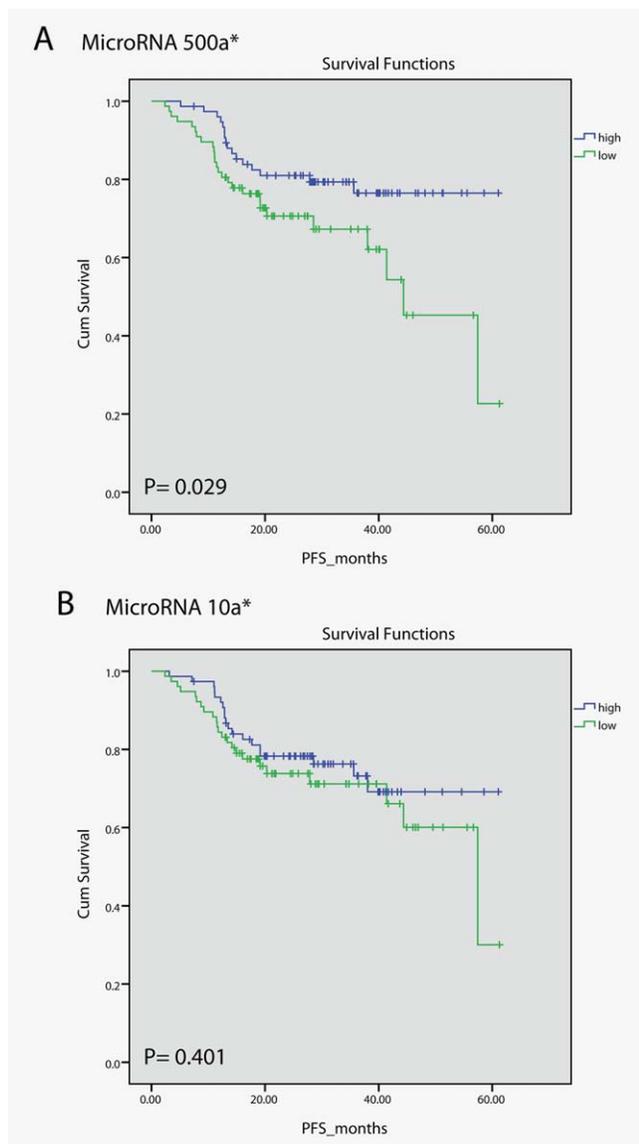
Figure 3. Kaplan–Meier curves of miR-500a* and miR-10a* from the RT-qPCR analysis. The microRNA expression values were dichotomized to high or low expression based on the median expression. (*a*) High expression of miR-500a* was significantly associated with better prognosis than low miR-500a* expression (*p* = 0.029). (*b*) The expression of miR-10a* did not reach significance in the validation analysis using RT-qPCR.

microRNAs with significantly different expression in the lung adenocarcinomas compared with normal lung tissue. Some of the microRNAs differentially expressed are known from other studies,[14] and some have not been previously identified. *EGFR*-mutated tumors exhibited significantly different microRNA-profiles compared with wt tumors. Moreover, the miR-500a* expression was associated with time to progression.

MicroRNAs differentially expressed between lung adenocarcinomas and normal lung tissue
The hierarchical clustering showed that lung adenocarcinomas and normal lung tissue samples had different microRNA

expression profiles, and 129 microRNAs were differentially expressed between the two groups. 95 microRNAs had lower expression in tumor tissue than in the normal lung tissue.

MiR-133a is the most downregulated microRNA in lung adenocarcinoma tissue compared with normal lung tissue in our study. The expression of this microRNA has been found to be altered in various types of cancer[22] but not in lung adenocarcinomas to our knowledge. MiR-133a was downregulated in a study of squamous cell lung carcinomas. When the expression was restored in SCC cell lines, this resulted in significant inhibition of cell proliferation, suggesting that miR-133a may function as a tumor suppressor.[23] MiR-9* is the microRNA most upregulated in our study, with a fold change of 29.1. Previous studies have not shown this microRNA to be differentially expressed between NSCLC and normal lung tissue, but it has been shown to be upregulated in cervical carcinomas.[24] MiR-9 has in most studies been found to be overexpressed in lung cancers,[14,25] and in our study miR-9 was upregulated 7.5 fold (Supporting Information Table S3).

Guan *et al.*[14] have reviewed the microRNA expression profiling-studies comparing lung cancer tissues with normal lung tissue, and found that some microRNAs are consistently reported. They performed a subgroup analysis on histological type, and in four studies that includes only adenocarcinoma patients, seven microRNAs were consistently reported. These microRNAs are differentially expressed and deregulated in the same direction as reported in the review, in our study. Of the 54 reported microRNAs differentially expressed between NSCLC (all histologies) and normal lung tissue, we identified 35. This may be due to different microRNA profiles in squamous cell carcinomas and adenocarcinomas of the lung.[26]

MiR-21 has been shown to have a higher expression in lung cancer tissue compared with normal lung tissue[11,27] and in other cancer types such as breast[28] and colon cancer.[10] Increased expression of miR-21 has been reported to be independently associated with reduced survival in NSCLC,[29,30] but it was not significant in the large International Adjuvant Lung Cancer Trial (IALT),[31] nor was it associated with time to progression in our study.

Previous studies have reported miR-126 as downregulated in NSCLC compared with normal lung tissue[14] and it is considered to be a tumor suppressor microRNA that inhibits tumor growth by preventing cancer cell proliferation, migration and invasion.[32] Our microarray results confirm the aberrant expression of miR-21 and miR-126, also validated in our independent validation cohort.

MicroRNAs differentially expressed between *EGFR*-mutated and *EGFR* wt lung adenocarcinomas
We identified 17 microRNAs differentially expressed between *EGFR*-mutated lung adenocarcinomas (*n* = 22) and *EGFR* wt lung adenocarcinomas (*n* = 130) (Table 2).

Not much is known about the microRNA expression in *EGFR*-mutated versus wt adenocarcinomas of the lung. Using microRNA microarrays, Dacic *et al.*[33] analyzed six lung adenocarcinomas (2 *EGFR*-positive, 2 *KRAS*-positive and 2

KRAS- and *EGFR*-negative) and found several microRNAs differentially expressed between the three groups. Four microRNAs were confirmed in a validation group of 18 samples using RT-qPCR. None of the four microRNAs differentially expressed (miR-155, miR-25, miR-495 and miR-7g) between the three groups were confirmed in our study. This may be due to their small sample size.

Seike *et al.*[27] identified 12 microRNAs differentially expressed between 6 *EGFR*-mutated tumors and 22 wildtype tumors. The tumors in their study had different histology, and all patients were never smokers. Some of the differentially expressed microRNAs between tumor tissue and normal lung tissue were validated in our study, but none of the 12 microRNAs that were differentially expressed between *EGFR*-mutated tumors and wt tumors.

MiR-452 is not previously known to be associated with lung adenocarcinomas. It has been shown to be altered in breast cancer.[34] The expression of miR-452 was altered in breast cancer cell lines after stimulating the cells with an epidermal growth factor, indicating that this microRNA may be part of the EGFR pathway.[35]

MicroRNA profiles associated with different clinicopathological features and the *KRAS*-mutational status

There were only a few microRNAs differentially expressed based on the patients' sex, smoking status, age and tumor stage. Only miR-516-5p was differentially expressed between male and female, and miR-3137 was differentially expressed between patients older or younger than 65 years. MiR-450 and miR-532-5p were differentially expressed between smokers/former smokers and never smokers. Seike's study[27] reports significantly fewer aberrantly expressed microRNAs in adenocarcinomas from never-smokers than in putatively smoking-induced tumors. This might be because microRNA genes often are located at fragile sites and regions with allelic imbalance, more commonly altered in smoking induced cancers.[36] The tumors harboring *KRAS* mutation had a lower expression of miR-371-5p and miR-564 and a higher expression of miR-100. Voortman and coworkers identified a borderline significant association between *KRAS* mutations and let-7a, which we could not confirm in our work.[31]

MicroRNA target prediction

We identified 10,197 mRNAs to be highly predicted or experimentally observed to be targeted by the 129 microRNAs differentially expressed between lung adenocarcinoma tissue and normal lung tissue. MiR-21 is shown to suppress *PTEN* in NSCLC tumors,[37] and miR-31 regulates *CDKN2A* in cell lines.[38] *MDM2* is predicted to be targeted by miR-138 and miR-30b*, both downregulated in our experiment.

We found 1,254 mRNA targets predicted or observed to be targeted by 6 of the 17 microRNAs that were differentially expressed between *EGFR*-mutated and *EGFR* wt tumors. MiR-184 has been observed to target *AKT2*,[20] which is an

important gene in the phosphatidylinositol 3-kinase (PI3K) pathway.[39] Mir-654-5p has been predicted to target *AKT3* that act downstream of *EGFR* to regulate many intracellular processes like cell survival, proliferation and growth. Both miR-30d* and 502-3p have been predict to target *CDK6*, an important regulator of cell-cycle progression and a known downstream molecule in the EGFR pathway.

MicroRNAs that can act as prognostic biomarkers

Previous studies have observed individual microRNAs or sets of microRNAs with prognostic information for NSCLC patients.[11,26,40]

In our study, high expression of miR-10a* and low expression of miR-500a* were most significantly related with shorter time to progression in the microarray analyses. MiR-500a* had a higher expression in *EGFR*-mutated tumors. Using IPA and pubmed search, we did not find any known mRNA targets for miR-10a* or miR-500a*. MiR-10a* has been previously described as altered in cervical carcinomas.[41] We did not manage to confirm the prognostic association of miR-10a* by RT-qPCR. This might be because the validation was done by TaqMan analyses is more quantitative and specific than the microarray analyses. MiR-500a* has to our knowledge not previously been associated with prognosis in NSCLC, and was confirmed to be associated with prognosis by RT-qPCR in the same cohort in our study.

Lung cancer mortality and poor prognosis are strongly associated with the predominant diagnosis of late-stage disease, and molecular biomarkers may be useful for early detection of the disease. Specific microRNAs have been found associated with prognosis in plasma samples from lung cancer patients.[42,43] MicroRNAs may also act as potential targets in cancer therapy.[44]

Previous studies of microRNA expression in lung cancer have focused on a limited fraction of the known human microRNAs. The increasing use of high-throughput technologies like microarrays and high-throughput sequencing has made it possible to study a larger number of microRNAs in a larger number of samples. Earlier studies have not separated the lung cancer tissue according to histologies that we know have a great impact on the microRNA expression.[26] We have focused on lung adenocarcinomas, which are increasing in incidence both in men and women. We have analyzed the mutation status of *EGFR* and *KRAS* of almost all tumors, and we have clinical information and follow-up information on all patients included.

Conclusion

Precise molecular mechanisms for the altered expression of microRNAs in lung cancer tumors are unclear. Abnormal expression may be caused by somatic genetic alterations and epigenetic mechanisms. Our study identified several novel microRNAs differentially expressed in adenocarcinomas of the lung compared with normal lung tissue, and the study also confirmed several microRNAs previously reported. Some of

these microRNAs may have the potential as biomarkers for diagnosis, as identified in tumor tissue, blood or sputum. Seventeen microRNAs were differentially expressed between *EGFR*-mutated and wt adenocarcinomas of the lung. This is important for classification of lung adenocarcinomas, and the differentially expressed microRNAs might be associated with genes in the EGFR pathway. We have also identified that miR-500a* can act as a potential prognostic marker for lung adenocarcinomas. As such, these findings may help us understand more of the biology of the adenocarcinomas of the lung and may be used as biomarkers, in classification, or as prognostic markers in the future.

References

1. Ferlay J, Shin HR, Bray F, *et al.* Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer* 2008;127:2893–917.
2. Sagerup CM, Smastuen M, Johannesen TB, *et al.* Sex-specific trends in lung cancer incidence and survival: a population study of 40,118 cases. *Thorax* 2011;66:301–7.
3. Devesa SS, Bray F, Vizcaino AP, *et al.* International lung cancer trends by histologic type: male:female differences diminishing and adenocarcinoma rates rising. *Int J Cancer* 2005;117:294–9.
4. Ding L, Getz G, Wheeler DA, *et al.* Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 2008;455:1069–75.
5. Saintigny P, Burger JA. Recent advances in non-small cell lung cancer biology and clinical management. *Discov Med* 2012;13:287–97.
6. Bartels CL, Tsongalis GJ. MicroRNAs: novel biomarkers for human cancer. *Ann Biol Clin (Paris)* 2010;68:263–72.
7. John B, Enright AJ, Aravin A, *et al.* Human MicroRNA targets. *PLoS Biol* 2004;2:e363.
8. Lewis BP, Burge CB, Bartel DP. Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. *Cell* 2005;120:15–20.
9. Iorio MV, Visone R, Di Leva G, *et al.* MicroRNA signatures in human ovarian cancer. *Cancer Res* 2007;67:8699–707.
10. Schetter AJ, Leung SY, Sohn JJ, *et al.* MicroRNA expression profiles associated with prognosis and therapeutic outcome in colon adenocarcinoma. *JAMA* 2008;299:425–36.
11. Yanaihara N, Caplen N, Bowman E, *et al.* Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* 2006;9:189–98.
12. Lu J, Getz G, Miska EA, *et al.* MicroRNA expression profiles classify human cancers. *Nature* 2005;435:834–8.
13. Vosa U, Vooder T, Kolde R, *et al.* Identification of miR-374a as a prognostic marker for survival in patients with early-stage nonsmall cell lung cancer. *Genes Chromosom Cancer* 2011;50:812–22.
14. Guan P, Yin Z, Li X, *et al.* Meta-analysis of human lung cancer microRNA expression profiling studies comparing cancer tissues with normal tissues. *J Exp Clin Cancer Res* 2012;31:54.
15. Helland A, Skaug HM, Kleinberg L, *et al.* EGFR gene alterations in a Norwegian cohort of lung cancer patients selected for surgery. *J Thorac Oncol* 2011;6:947–50.
16. Hamfjord J, Stangeland AM, Skrede ML, *et al.* Wobble-enhanced ARMS method for detection of KRAS and BRAF mutations. *Diagn Mol Pathol* 2011;20:158–65.

17. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 2001;25:402–8.
18. Dysvik B, Jonassen I. J-Express: exploring gene expression data using Java. *Bioinformatics* 2001;17:369–70.
19. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. *Proc Natl Acad Sci USA* 2001;98:5116–21.
20. Foley NH, Bray IM, Tivnan A, *et al.* MicroRNA-184 inhibits neuroblastoma cell survival through targeting the serine/threonine kinase AKT2. *Mol Cancer* 2010;9:83.
21. Nakamoto M, Jin P, O'Donnell WT, *et al.* Physiological identification of human transcripts translationally regulated by a specific microRNA. *Hum Mol Genet* 2005;14:3813–21.
22. Nohata N, Hanazawa T, Enokida H, *et al.* microRNA-1/133a and microRNA-206/133b clusters: dysregulation and functional roles in human cancers. *Oncotarget* 2012;3:9–21.
23. Moriya Y, Nohata N, Kinoshita T, *et al.* Tumor suppressive microRNA-133a regulates novel molecular networks in lung squamous cell carcinoma. *J Hum Genet* 2012;57:38–45.
24. Lee JW, Choi CH, Choi JJ, *et al.* Altered MicroRNA expression in cervical carcinomas. *Clin Cancer Res* 2008;14:2535–42.
25. Xu T, Liu X, Han L, *et al.* Up-regulation of miR-9 expression as a poor prognostic biomarker in patients with non-small cell lung cancer. *Clin Transl Oncol* 2013. DOI 10.1007/s12094-013-1106-1.
26. Landi MT, Zhao Y, Rotunno M, *et al.* MicroRNA expression differentiates histology and predicts survival of lung cancer. *Clin Cancer Res* 2010;16:430–41.
27. 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;106:12085–90.
28. Yan LX, Huang XF, Shao Q, *et al.* MicroRNA miR-21 overexpression in human breast cancer is associated with advanced clinical stage, lymph node metastasis and patient poor prognosis. *RNA* 2008;14:2348–60.
29. Markou A, Tsaroucha EG, Kaklamanis L, *et al.* Prognostic value of mature microRNA-21 and microRNA-205 overexpression in non-small cell lung cancer by quantitative real-time RT-PCR. *Clin Chem* 2008;54:1696–704.
30. Gao W, Shen H, Liu L, *et al.* MiR-21 overexpression in human primary squamous cell lung carcinoma is associated with poor patient prognosis. *J Cancer Res Clin Oncol* 2011;137:557–66.

31. Voortman J, Goto A, Mendiboure J, *et al.* MicroRNA expression and clinical outcomes in patients treated with adjuvant chemotherapy after complete resection of non-small cell lung carcinoma. *Cancer Res* 2010;70:8288–98.
32. Meister J, Schmidt MH. miR-126 and miR-126*: new players in cancer. *Sci World J* 2010;10:2090–100.
33. Dacic S, Kelly L, Shuai Y, *et al.* miRNA expression profiling of lung adenocarcinomas: correlation with mutational status. *Mod Pathol* 2010; 23:1577–82.
34. Volinia S, Galasso M, Sana ME, *et al.* Breast cancer signatures for invasiveness and prognosis defined by deep sequencing of microRNA. *Proc Natl Acad Sci USA* 2012;109:3024–9.
35. Avraham R, Sas-Chen A, Manor O, *et al.* EGF decreases the abundance of microRNAs that restrain oncogenic transcription factors. *Sci Signal* 2010;3:ra43.
36. Momi N, Kaur S, Rachagani S, *et al.* Smoking and microRNA dysregulation: a cancerous combination. *Trends Mol Med* 2014; 20:36–47.
37. Zhang JG, Wang JJ, Zhao F, *et al.* MicroRNA-21 (miR-21) represses tumor suppressor PTEN and promotes growth and invasion in non-small cell lung cancer (NSCLC). *Clin Chim Acta* 2010;411:846–52.
38. Malhas A, Saunders NJ, Vaux DJ. The nuclear envelope can control gene expression and cell cycle progression via miRNA regulation. *Cell Cycle* 2010;9:531–9.
39. Vivanco I, Sawyers CL. The phosphatidylinositol 3-Kinase AKT pathway in human cancer. *Nat Rev Cancer* 2002;2:489–501.
40. Lu Y, Govindan R, Wang L, *et al.* MicroRNA profiling and prediction of recurrence/relapse-free survival in stage I lung cancer. *Carcinogenesis* 2012;33:1046–55.
41. Li JH, Xiao X, Zhang YN, *et al.* MicroRNA miR-886-5p inhibits apoptosis by down-regulating Bax expression in human cervical carcinoma cells. *Gynecol Oncol* 2011;120:145–51.
42. Shen Y, Tang D, Yao R, *et al.* microRNA expression profiles associated with survival, disease progression and response to gefitinib in completely resected non-small-cell lung cancer with EGFR mutation. *Med Oncol* 2013;30:750.
43. Zhang H, Su Y, Xu F, *et al.* Circulating microRNAs in relation to EGFR status and survival of lung adenocarcinoma in female non-smokers. *PLoS One* 2013;8:e81408.
44. Cho WC. MicroRNAs: potential biomarkers for cancer diagnosis, prognosis and targets for therapy. *Int J Biochem Cell Biol* 2010;42:1273–81.

Cancer Genetics