

# Identification of MiR-374a as a Prognostic Marker for Survival in Patients with Early-Stage Nonsmall Cell Lung Cancer

Urmo Vösa,<sup>1</sup> Tõnu Vooder,<sup>2</sup> Raivo Kolde,<sup>3</sup> Krista Fischer,<sup>6</sup> Kristjan Väik,<sup>1</sup> Neeme Tõnisson,<sup>1,4</sup> Retlav Roosipuu,<sup>5</sup> Jaak Vilo,<sup>3</sup> Andres Metspalu,<sup>1,6</sup> and Tarmo Annilo<sup>1\*</sup>

<sup>1</sup>Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Riia 23, 51010 Tartu, Estonia

<sup>2</sup>Clinic of Cardiovascular and Thoracic Surgery of Tartu University Hospital, Puusepa 8, 51014 Tartu, Estonia

<sup>3</sup>Institute of Computer Science and Center of Translational Genomics, University of Tartu, Liivi 2, 50409 Tartu, Estonia

<sup>4</sup>Tartu University Hospital, Puusepa 8, 51014 Tartu, Estonia

<sup>5</sup>Department of Pathology, Hospital of Tartu University, Puusepa 8, 51014 Tartu, Estonia

<sup>6</sup>Estonian Genome Center and Center of Translational Genomics, University of Tartu, Tiigi 6Ib, 50410 Tartu, Estonia

Lung cancer is one of the deadliest types of cancer proven by the poor survival and high relapse rates after surgery. Recently discovered microRNAs (miRNAs), small noncoding RNA molecules, play a crucial role in modulating gene expression networks and are directly involved in the progression of a number of human cancers. In this study, we analyzed the expression profile of 858 miRNAs in 38 Estonian nonsmall cell lung cancer (NSCLC) samples (Stage I and II) and 27 adjacent nontumorous tissue samples using Illumina miRNA arrays. We found that 39 miRNAs were up-regulated and 33 down-regulated significantly in tumors compared with normal lung tissue. We observed aberrant expression of several well-characterized tumorigenesis-related miRNAs, as well as a number of miRNAs whose function is currently unknown. We show that low expression of miR-374a in early-stage NSCLC is associated with poor patient survival. The combinatorial effect of the up- and down-regulated miRNAs is predicted to most significantly affect pathways associated with cell migration, differentiation and growth, and several signaling pathways that contribute to tumorigenesis. In conclusion, our results demonstrate that expression of miR-374a at early stages of NSCLC progression can serve as a prognostic marker for patient risk stratification and may be a promising therapeutic target for the treatment of lung cancer. © 2011 Wiley-Liss, Inc.

## INTRODUCTION

Lung cancer is characterized by a low survival and high relapse rate after surgery and is the leading cause of cancer death (Jemal et al., 2010). The most common type is nonsmall cell lung cancer (NSCLC), which accounts for approximately 75% of lung cancers and has two major histological subtypes: squamous cell carcinoma (SCC) and adenocarcinoma (AD). Poor survival of lung cancer patients is largely due to the lack of reliable early diagnostic markers, often resulting in diagnosis of tumors at advanced stages. Recent technological innovations have enabled identification of novel molecular classifiers on the genomic scale. Combining biomarker and clinical data could be helpful for early tumor detection and lead to a more precise prediction of both prognosis and therapeutic response.

MicroRNAs (miRNAs) are small noncoding RNA molecules that interact with partially complementary sequences in the 3' UTR of the target mRNA and modulate gene expression at the post-transcriptional level. Such fine-tuning of the protein repertoire of the cell in numerous molec-

ular networks controls several central biological processes in many different species (Fabian et al., 2010; Ryan et al., 2010).

Aberrant regulation of miRNAs has been associated with a number of human diseases, including many cancers (excellent reviews include Croce, 2009; Ryan et al., 2010; Schaefer et al., 2010). In tumorigenesis, miRNAs may function as tumor suppressors (Chen et al., 2010) or oncogenes (Pineau et al., 2010). One of the first miRNAs to be associated with lung cancer

Additional Supporting Information may be found in the online version of this article.

Supported by Targeted Financing from the Estonian Ministry of Education and Research, Grant number: SF0180142s08; Estonian Science Foundation, Grant numbers: ETF6465 and ETF7473; EU FP7 ECOGENE, Grant number: #205419 and OPENGENE #245536, EBC; European Union via a European Regional Development Fund to the Centre of Excellence in Genomics, Estonian Biocentre and University of Tartu.

\*Correspondence to: Tarmo Annilo, Department of Biotechnology, Institute of Molecular and Cell Biology, University of Tartu, Riia 23, 51010 Tartu, Estonia. E-mail: tarmo.annilo@ut.ee

Received 30 November 2010; Accepted 9 June 2011

DOI 10.1002/gcc.20902

Published online 11 July 2011 in Wiley Online Library (wileyonlinelibrary.com).

pathogenesis was *let-7* (Takamizawa et al., 2004), an miRNA that was first discovered in *Caenorhabditis elegans* (Reinhart et al., 2000). It was later shown that *let-7* is an important tumor suppressor, controlling the expression of cellular oncogenes, such as *RAS* (Kumar et al., 2008). Large-scale miRNA expression profiling of tumors and developing tissue has identified several additional miRNA candidates that may contribute to tumorigenesis (Calin et al., 2004; Miska et al., 2004; Thomson et al., 2004). Furthermore, it appears that miRNAs can be used to differentiate histological subtypes and predict survival of cancer patients. For example, a profile of 34 miRNAs that clearly differentiate SCC and AD has been identified (Landi et al., 2010). This difference in miRNA expression was most significant in the earlier stages of disease and less significant in the late-stage tumors, indicating the heterogeneity and complicated histological nature of late-stage tumors. The capacity for miRNAs to regulate molecular networks that contribute to carcinogenesis and the remarkable stability of miRNAs in body fluids suggests that the development of clinically relevant miRNA-based biomarkers for disease detection and therapeutic monitoring is a feasible task (Mitchell et al., 2008).

The human genome is estimated to contain a little over a thousand different miRNA genes (miRBase Release 16 reports 1,048 human miRNAs as of September 2010; www.mirbase.org). Unfortunately, as the number of known miRNAs has dramatically increased during the last few years, earlier studies investigated only a very limited fraction of the total human miRNA repertoire. In addition, mature miRNAs may be subjected to post-transcriptional modifications (Ebhardt et al., 2009), further complicating the analysis. A number of technical issues, including population origin, method of miRNA extraction, quantitation, data analysis, and endogenous control choice for qRT-PCR analysis, can also lead to inconsistent results. Therefore, extensive replication and validation using independent study groups, more advanced technologies, and representative miRNA sets are necessary to prove the clinical relevance of miRNAs as biomarkers.

In this study, we have performed miRNA expression analysis of Stage I and II NSCLC samples with the aim of identifying miRNA alterations that are likely to be causative in tumorigenesis rather than later-stage “passenger” aberrations. The samples were collected from an ethnically homogeneous Estonian population of European descent (Nelis et al., 2009) and profiled

TABLE 1. Clinical and Pathological Characteristics of the Patient Samples

Characteristic	SCC	AD	Normal lung <sup>a</sup>
Number of samples	18	20	27
Sex			
Males	18	14	23
Females	0	6	4
Stage			
Ia	2	12	7
Ib	14	7	17
IIa	0	0	0
IIb	2	1	3
Smoker			
Yes	17	15	24
No	0	4	3
Unknown	1	1	
Age			
≤ 63 years	7	9	11
> 63 years	11	11	16
Range	44–79	41–79	41–77
Median	67	64	67
Life status			
Alive	9	16	20
Deceased	9	4	7

<sup>a</sup>Normal lung, cancer-free tissue adjacent to tumor from NSCLC patients.

Abbreviations: SCC, squamous cell carcinoma; AD, adenocarcinoma.

using Illumina microarrays. We also analyzed the prognostic value of miRNA expression analysis and found that miR-374a expression is associated with patient survival.

## MATERIALS AND METHODS

### Patients and Collection of Tissue Samples

NSCLC and cancer-free tissue samples were collected from patients who had undergone curative resection and histologically characterized in Tartu University Lung Hospital (Estonia) from 2002 to 2008. The study was approved by the Ethics Committee on Human Research of Tartu University and written informed consent was obtained from all patients. This study incorporated 38 tumor and 27 adjacent normal lung samples (including 24 paired cases) from patients with Stage I or II NSCLC, who had not received radio- or chemotherapy before surgery. After surgery, six patients were treated with adjuvant chemotherapy. These data were enrolled into the multivariate Cox regression analyses as a covariate. Staging was based on the International System for Staging Lung Cancer (Mountain, 2000). The clinical and pathological characteristics of the study subjects are shown in Table 1.

### RNA Isolation and miRNA Expression Profiling Using Illumina BeadChips

Total RNA was extracted from 25 mg of snap-frozen tumor or normal lung tissue using Trizol Reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions, and further purified using the RNeasy Mini Kit (Qiagen, Valencia, CA). RNA quality was assessed using the Agilent 2100 Bioanalyzer and concentration was measured using Nanodrop 1000 (Thermo Fisher Scientific, Waltham, MA). Total RNA (400 ng) was hybridized to the Illumina MicroRNA Profiling BeadChip, containing 858 mature human miRNA probes and 287 hypothetical small RNA probes according to the standard protocol provided by Illumina at the Estonian Biocentre Core Facility, Tartu, Estonia.

### Real-Time PCR Quantification

Microarray results were validated by qRT-PCR analysis using TaqMan miRNA assays (Applied Biosystems, Beverly, MA). First, we analyzed the expression of four miRNAs in eight pairs of tumor and cancer-free lung samples that were included in the Illumina microarray sample set. Second, we analyzed a group of eight additional sample pairs that were not included in the microarray sample set. Reverse transcription with miRNA-specific stem-loop primers was performed using 40 ng of total RNA and the First Strand cDNA Synthesis Kit (Fermentas Inc., Glen Burnie, MD) according to the manufacturer's instructions. Reactions were incubated at 16°C for 30 min, 42°C for 30 min, and 85°C for 5 min. Real-time PCR reactions were performed in triplicate using the ABI Prism 7900 instrument (Applied Biosystems) using TaqMan assays (Applied Biosystems, Part Number 4427975). Cycling conditions were 95°C for 15 min, followed by 40 cycles of 95°C for 15 sec and 60°C for 1 min. The expression of RNU48, a small nucleolar RNA, or the geometric mean of two highly and consistently expressed miRNAs, miR-16 and miR-26b (Vandesompele et al., 2002; Davoren et al., 2008), was used as an endogenous expression reference. The relative expression of each miRNA was calculated according to  $2^{-\Delta\Delta C_t}$  method.

### Statistical Analysis

Array data were exported from Illumina GenomeStudio (v2009.2) and analyzed using *lumi*, a Bioconductor package specifically

designed for processing and analysis of the Illumina microarray data written in R (Du et al., 2008). Hypothetical small RNA capture probes were excluded from the subsequent analyses. The technical quality of the microarray chips was evaluated using the package *arrayQualityMetrics*. In addition, the probe control data, including controls for poly(A)-polymerase, annealing, extension, hybridization, internal mismatch, contamination, and miRNA intensity, were examined for each sample. Raw expression data was background subtracted using negative control probe information ("bgAdjust" method in *lumi*) and an offset (minus minimum value plus one) was added to all expression values ("forcePositive" method in *lumi*). For further analysis, only the probes that showed statistically significant expression over the background level (detection *P*-value < 0.01) in more than half of the samples were selected, resulting in 506 miRNAs that were included in further analysis. Data was then  $\log_2$ -transformed and quantile-normalized using the corresponding functions from *lumi*. Linear models from *limma* were used to identify differentially expressed miRNAs using a *P*-value cutoff of 0.01. The Benjamini-Hochberg method was used to correct the *P*-values for false discovery rate (FDR). For pairwise analysis, we used a subset of samples where data from both tumor and normal lung tissue from the same individual was available ( $n = 24$ ).

### miRNA Target Prediction and Pathway Analyses

Enrichment analyses for KEGG pathways were performed using the Bioconductor package, *GOstats*. First, the miRNA targets were predicted using GeneMir. To get a more comprehensive look at the impact of miRNA alterations on the cell physiology, we analyzed the combinatorial effect of the 10 most up- or down-regulated miRNAs, together, instead of analyzing the influence of just a single miRNA. A gene was considered a target of a particular miRNA if predicted so by at least three out of five target prediction algorithms implemented in GeneMir (Pictar 4-way, PITA, TargetscanS, miRanda, and DIANA-microT). For miRNA expression data, the gene universe was defined as all predicted miRNA target genes for all miRNAs acquired using the GeneMir program, and the input list consisted of the predicted targets of the 10 most up- or down-regulated miRNAs. For mRNA expression data, the gene universe was defined as all genes

present on the mRNA array and the input list consisted of up- or down-regulated genes acquired from our previous study (Välk et al., 2010). Predicted miRNA target lists and gene lists from mRNA analyses were compared and analyzed using Fisher's exact test to determine significance.

### Survival Analysis

Survival time was calculated from the date of resection to the date of death or last follow-up. Two patients had a very short survival time after resection (<15 days), indicating a possible non-cancer related death, and were excluded from analysis. Cox proportional regression model was then applied to tumor samples ( $n = 36$ ) to identify miRNAs whose expression was significantly associated with survival. Model was used in the permutation test, where survival time and status (censoring indicator) were assigned randomly to each patient in 10,000 permutations to get permuted  $P$ -values (permutation  $P$ -value cutoff = 0.01).

Smoking status in pack-years, histological type (SCC or AD), tumor stage (Ia, Ib, or IIb), patient sex, age, and chemotherapy after surgery were used as covariates in subsequent Cox regression analyses to assess the effect of those potential confounders ( $P$ -value cutoff = 0.05). Identified genes which showed covariate-independent association were used in Kaplan-Meier estimation and log-rank test, where high-expression and low-expression groups were defined based on the miRNA median expression value.

## RESULTS

### miRNA Expression Profile in Estonian NSCLC Samples

We analyzed miRNA expression profiles in 38 NSCLC and 27 cancer-free control samples (including 24 paired cases, Table 1) using Illumina miRNA BeadChips. Thirty-nine miRNAs were found to be up-regulated and 33 down-regulated in tumors compared with normal lung tissue ( $P < 0.01$ , Fig. 1A). The 30 most significantly altered miRNAs are shown in Table 2, and the full list is given in Supporting Information Table 1. Application of an expression change cutoff value of two-fold as a proxy for biological significance reduced this number to 31 up- and 29 down-regulated miRNAs (Fig. 1A). We observed that the expression of several previously characterized tumorigen-

esis-associated miRNAs was altered in tumor samples; miR-9, miR-205, and miR-196a were up-regulated 15.5-, 12.3-, and 8.6-fold, respectively, and miR-101 and miR-206 were down-regulated 4.1- and 11.1-fold, respectively. Pairwise comparison demonstrated consistent expression changes in most of the normal and tumor sample pairs (Fig. 1B). In addition, our analysis revealed significant dysregulation of a number of miRNAs with no previous functional connection to human cancers, including up-regulation of miR-708 and miR-941 (5.5- and 4.3-fold, respectively) and down-regulation of miR-1273 (14.3-fold).

### Validation of the Microarray Data

First, the Illumina BeadChip results for four miRNAs were validated by real-time quantitative RT-PCR (qRT-PCR) in a subset of eight pairs of samples used in the microarray (Fig. 1C). The expression levels of miR-9, miR-149, miR-196a, and miR-205 were significantly higher in tumor samples compared with normal lung, confirming the Illumina BeadChip data. Second, we used eight independent sample pairs not included in the array analysis to validate the expression of 10 miRNAs (Fig. 1D, Supporting Information Fig. 1). For validation, the most aberrantly regulated miRNAs for which commercial assays were available were selected. All but one of the investigated miRNAs, miR-1273, were detectable by qRT-PCR in our samples.

To investigate the effect of the endogenous reference gene(s) chosen for analysis on the qRT-PCR results, we used two different endogenous references, RNU48 and the geometric mean of two miRNAs (miR-16 and miR-26b). Results show relatively high and consistent expression of miR-16 in both tumor and cancer-free lung samples with a median  $Ct$  of 15, while miR-26b was moderately abundant (median  $Ct = 19$ ). The correlation between the two normalization methods was statistically significant ( $R = 0.59$ – $0.94$ ,  $P < 0.05$ ) for six investigated miRNAs (Supporting Information Fig. 2), whereas for three miRNAs the correlation was not significant. In the majority of cases, the expression changes determined by the two different normalization methods and array analysis showed the same directionality (Fig. 1D).

### Survival Analysis

To investigate the association between miRNA expression levels and patient survival, the

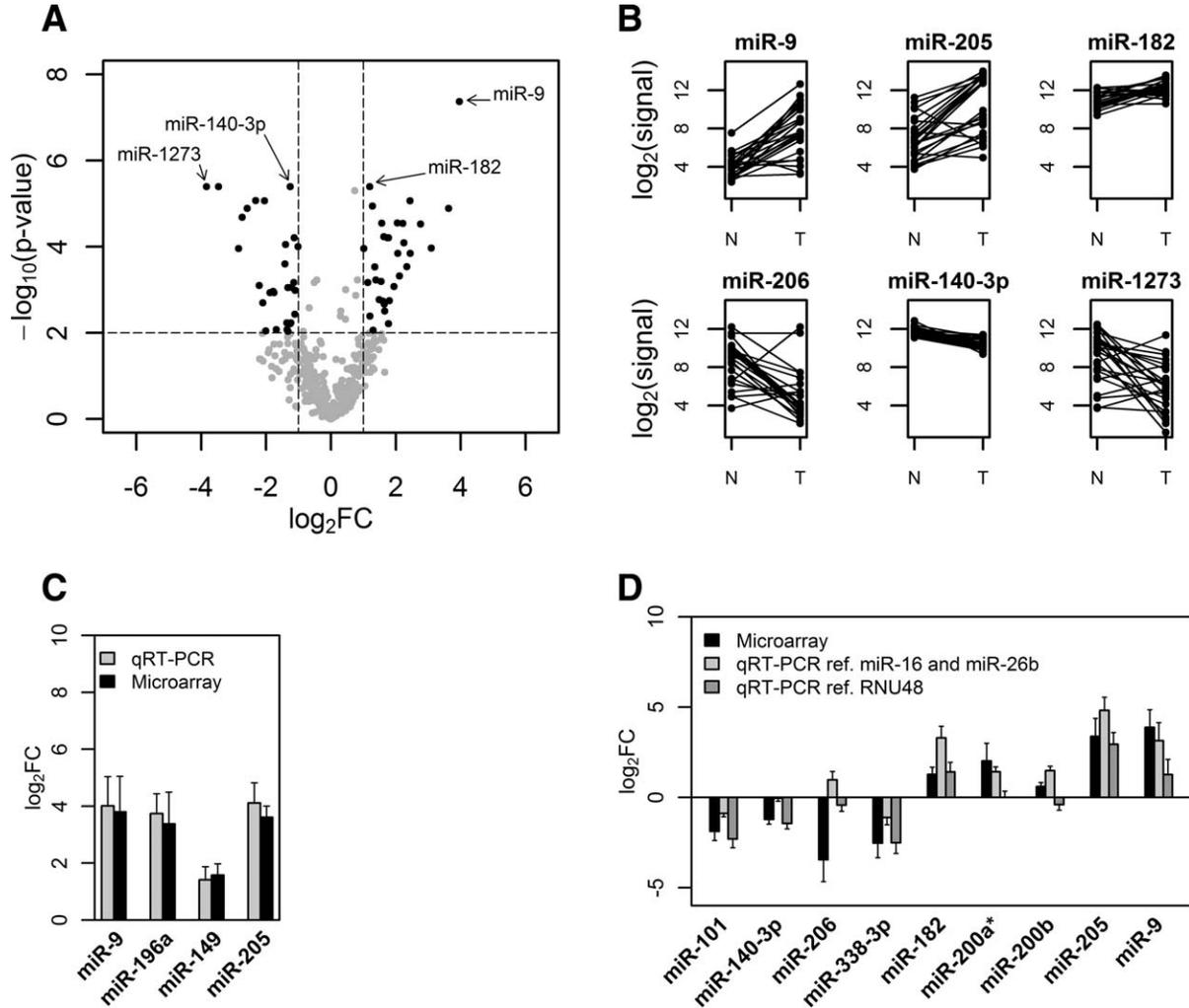


Figure 1. Altered expression of miRNA genes in NSCLC. A. Volcano plot depicting the results of Illumina microarray analysis comparing tumor and normal lung tissue samples. X-axis shows the expression level difference on the logarithmic scale ( $\log_2FC$ ), and the y-axis shows the statistical significance of the change [negative  $\log_{10}(P\text{-value})$ ]. Vertical dashed line indicates two-fold expression difference used as the arbitrary cut-off for biological impact. The level of statistical significance ( $P = 0.01$ ) after correction for multiple testing is marked as a horizontal dashed line. B. miRNA expression levels in 24 normal lung (N) and tumor tissue (T) sample pairs from the same patient. The change in expression is more pronounced for miR-

9 and miR-205 than for miR-182 and miR-140-3p. C. Validation of microarray results using qRT-PCR. Grey bars show the average expression difference between tumor and normal lung according to qRT-PCR. The geometric mean of miR-16 and miR-26b was used as an endogenous reference. Black bars show the expression difference according to Illumina miRNA microarray analysis. Data is presented as the SEM as calculated by  $\log_2$  value of expression change ( $\log_2FC$ ). D. Independent validation of microarray results by qRT-PCR of eight sample pairs not profiled on the array. Normalization was performed either using the geometric mean of miR-16 and miR-26b or RNU48 expression as an endogenous reference. Error bars indicate SEM.

univariate Cox proportional hazards regression model was fitted to the data for each miRNA. Eight miRNAs showed association with overall survival (permuted  $P$ -value  $< 0.01$ , Supporting Information Table 2).

To identify miRNAs which can serve as independent prognostic factors, multivariate Cox regression analysis with age, sex, smoking status, tumor subtype, stage, and postoperative chemotherapy (yes/no) as covariates was performed. We identified one miRNA whose expression had a significant effect on survival according to the

multivariate Cox regression model (HR = 0.353,  $P = 0.008$ , Supporting Information Table 3), miR-374a. We next defined the groups of tumor samples with high or low miRNA-374a expression based on the median expression value. Kaplan-Meier estimation and log-rank test showed significant difference in survival times between the high- and low-expression groups (Fig. 2). High expression of miR-374a increased overall survival, whereas low expression was associated with a significantly poorer prognosis ( $P = 0.018$ ).

TABLE 2. The 30 Most Differentially Expressed miRNAs in NSCLC Samples

	miRNA	P-value	FC	Host gene(s) <sup>a</sup>	Location	
Up-regulated	miR-9	4.30 e-08	15.5	C1orf61 [+]	1q22	
				LOC64532 [-]	5q14.3	
					Intergenic	15q26.1
	miR-182	4.06 e-06	2.3	Intergenic	7q32.2	
	miR-200a*	8.67 e-06	5.4	Intergenic	1p36.33	
	miR-151:9.1 <sup>b</sup>	1.15 e-05	2.4	PTK2 [+]	8q24.3	
	miR-205	1.29 e-05	12.3	LOC642587 [+]	1q32.2	
	miR-183	2.82 e-05	4.1	Intergenic	7q32.2	
	miR-130b*	2.87 e-05	3.0	Intergenic	22q11.21	
	miR-149	2.90 e-05	4.6	GPC1 [+]	2q37.3	
	miR-193b	2.98 e-05	6.8	Intergenic	16p13.12	
	miR-339-5p	5.86 e-05	3.1	C7orf50 [+]	7p22.3	
	miR-196b	6.19 e-05	3.4	Intergenic	7p15.2	
	miR-224	6.19 e-05	3.4	GABRe [+]	Xq28	
	miR-31	8.20 e-05	4.8	LOC554202 [+]	9p21.3	
	miR-196a	1.08 e-04	8.6	Intergenic	17q21.32	
				Intergenic	12q13.13	
		miR-423-3p	1.10 e-04	2.0	CCDC55 [+]	17q11.2
		miR-708	1.44 e-04	5.5	ODZ4 [+]	11q14.1
		miR-106b*	1.44 e-04	4.2	MCM7 [+]	7q22.1
	miR-210	2.92 e-04	5.1	Intergenic	11p15.5	
Down-regulated	miR-1273	4.06 e-06	14.3	RGS22 [+]	8q22.2	
	miR-206	4.06 e-06	11.1	Intergenic	6p12.2	
	miR-140-3p	4.06 e-06	2.4	VWPP2 [+]	16q22.1	
	miR-338-3p	8.51 e-06	5.0	AATK [+]	17q25.3	
	miR-101	3.5 e-05	4.1	Intergenic	1p31.3	
				RCL1 [+]	9p24.1	
	miR-144:9.1 <sup>#</sup>	1.29 e-05	6.0	Intergenic	17q11.2	
	miR-1285	2.08 e-05	6.7	KRIT1	17q11.2	
	miR-130a	6.19 e-05	2.2	Intergenic	11q12.1	
	miR-486-5p	8.97 e-05	2.6	ANK1 [+]	8p11.21	
	miR-24-2*	9.94 e-05	2.0	Intergenic	19p13.12	
	miR-144*	1.10 e-04	7.2	Intergenic	7q21.2	
	miR-30a	2.51 e-04	2.7	C6orf155	6q13	

<sup>a</sup>[+] indicates that the miRNA and its host gene are transcribed in the same direction.

<sup>b</sup>9.1, the given probe is specific to miRBase release 9.1 version of the miRNA.

Identical mature forms of miR-9, miR-196a, and miR-101 are encoded by more than one gene.

Abbreviation: FC, fold change.

### The Impact of miRNA Alterations on Regulatory Pathways

We investigated the potential consequences of altered miRNA expression on cellular regulatory pathways by enrichment analyses using the KEGG PATHWAY Database. Stringently predicted targets of the 10 most up-regulated (1,750 predicted targets) and down-regulated (1,909 predicted targets) miRNAs were used. The targets of aberrantly regulated miRNAs in our model were mostly associated with processes linked to cell mobility, differentiation, and proliferation, including axon guidance, and regulation of the actin cytoskeleton and focal adhesion, as well as other signaling and tumor-related pathways (Fig. 3). For comparison, we performed the same analyses with lists of significantly deregulated protein-coding genes acquired from our previous gene expression

study (575 down-regulated and 395 up-regulated genes; Valk et al., 2010). Analysis of data showed that up-regulated genes most strongly regulate cell proliferation-associated processes (cell cycle, p53 signaling, and DNA replication), while down-regulated genes primarily modulate pathways involved in immunity and the inflammatory response (Supporting Information Figure 3). We also performed target enrichment analysis for miR-374a and found that its targets were most strongly associated with cancer-related pathways (687 targets, FDR-corrected *P*-value = 0.024).

### Concordance Between miRNA and mRNA Expression Data

To test whether deregulated genes from our previous mRNA profiling study (Valk et al., 2010) are over-represented among predicted targets of

most deregulated miRNAs, we compared these gene lists using Fisher's exact test (Table 3). Genes showing down-regulation in our previous gene expression study were significantly over-represented among targets of ten most up-regulated miRNAs identified in this study [ $P = 0.021$ , odds ratio (OR) = 1.37, 95% confidence interval (CI) 1.04–1.78]. Interestingly, genes showing up-regulation in our mRNA profiling study were rather under-represented among the targets of ten most down-regulated miRNAs ( $P = 0.0004$ , OR = 0.49, 95% CI 0.30–0.75). Comparisons of two up-regulated and two down-regulated gene lists did not yield any statistically significant results ( $P > 0.05$ ).

## DISCUSSION

We have investigated miRNA expression alterations in early-stage NSCLC using Illumina microar-

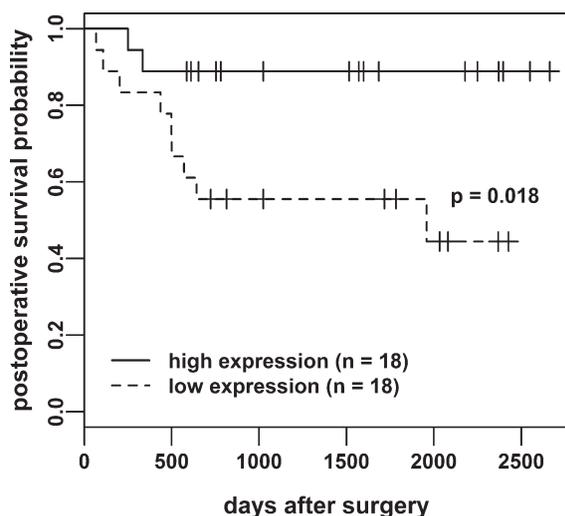


Figure 2. Kaplan-Meier plot showing the overall survival of patients according to the expression level of miR-374a. Patients were stratified into high- and low-expression groups based on the miRNA median expression value; high expression is considered greater than the median value of 12.16. The log-rank test  $P$ -value = 0.018. High expression of miR-374a is associated with better survival.

rays containing 858 miRNA probes, which is one of the most up-to-date human miRNA profiling arrays at this time. The expression profile of tumor samples in our study was clearly different from normal lung tissue; several of the most dysregulated miRNAs, including miR-9, miR-205, and miR-206, showed significant and consistent expression changes in tumor vs. normal lung pairwise analysis (Fig. 1B). Previous studies have demonstrated that miR-9 induces increased cell migration and invasiveness (Ma et al., 2010). Down-regulation of miR-206 in our samples is consistent with its functional role as an apoptosis activator and suppressor of tumor cell migration (Song et al., 2009).

Comparison of our results with previously published studies investigating miRNA expression in NSCLC tumor tissue versus normal lung demonstrates that although different studies have revealed largely nonoverlapping lists of altered miRNAs, there is a small set of miRNAs that are reported to be dysregulated in most of the studies (Table 4). The number of miRNAs screened, population origin, sample size, and the detection methods vary between these studies, but seven papers out of 10 (including this study) have reported down-regulation of miR-30a and miR-126. miR-30a is a regulator of liver development and function (Hand et al., 2009). miR-126 is crucial for angiogenesis and vascular maintenance (Fish et al., 2008) and has been reported to have tumor suppressor activity (Liu et al., 2009). The expression change of miR-126 in our study was not large (2.3 times, Supporting Information Table 1) but very consistent in pairwise analysis, indicating that miR-126 is a good candidate for further functional experiments. miR-126 is located within the intron of *EGFL7* (epidermal growth factor like-7), which positively regulates angiogenesis (Fish et al., 2008; Kuhnert et al., 2008).

Six studies (including this study) have reported up-regulation of miR-17 and miR-210. miR-17 is a member of the mir-17-92 oncogenic cluster (Dews

TABLE 3. Frequency Table Demonstrating the Overrepresentation of Down-Regulated Genes From the Gene Expression Study Among Predicted Targets of the 10 Most Up-Regulated miRNAs (A) and Underrepresentation of Up-Regulated Genes Among Predicted Targets of the 10 Most Down-Regulated miRNAs (B)

	Down-regulated genes	Nondifferentially expressed genes	Odds ratio	95% CI	$P$ -value
<b>A</b>					
Predicted targets	68 (13%)	1,418 (9.8%)	1.37	1.04–1.78	0.021
Other genes	455 (87%)	12,960 (90.2%)			
	Up-regulated genes	Nondifferentially expressed genes			
<b>B</b>					
Predicted targets	23 (6.3%)	1,767 (12.2%)	0.49	0.30–0.75	0.0004
Other genes	340 (93.7%)	12,771 (87.8%)			

Significance was determined by Fisher's exact test.

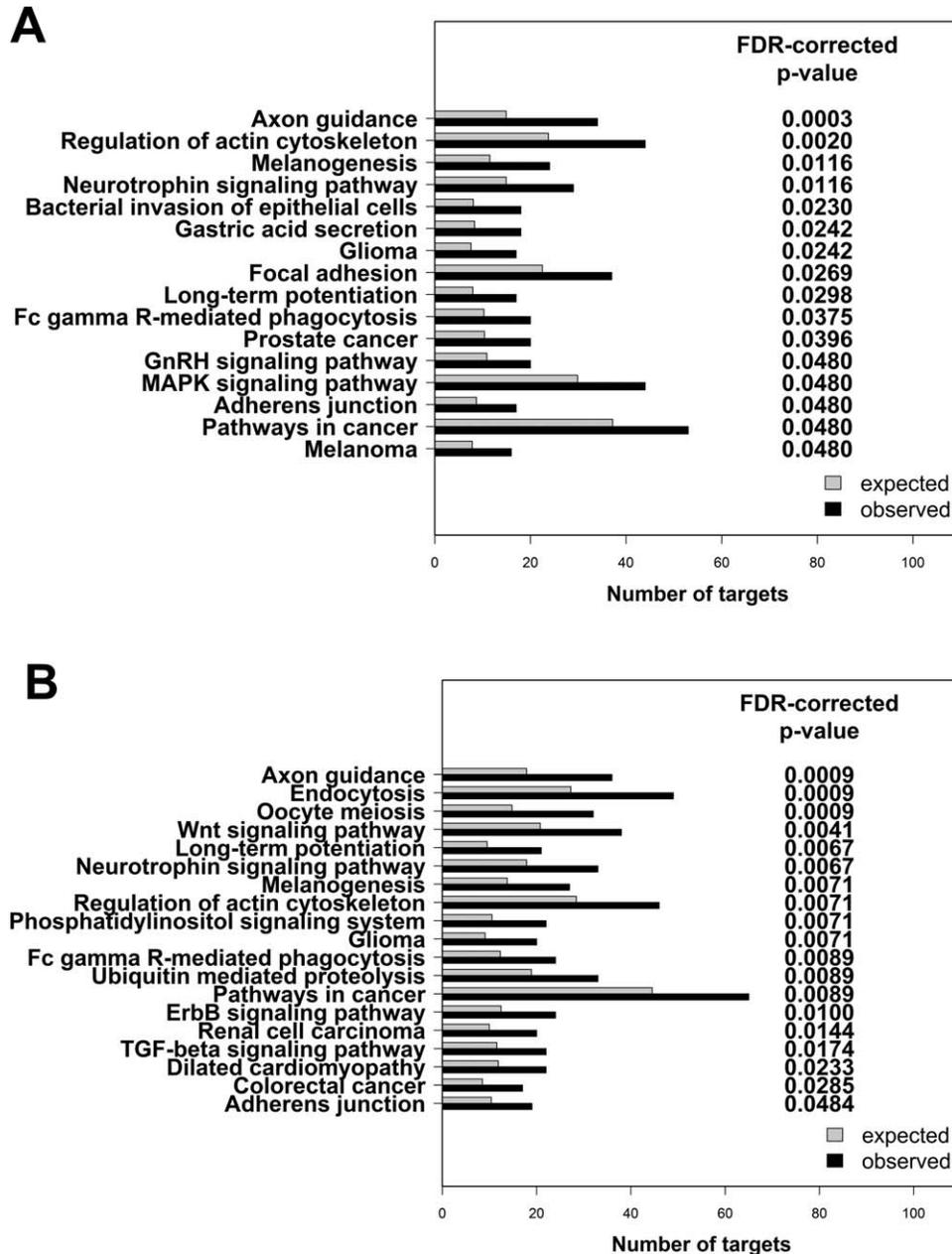


Figure 3. Identified KEGG pathways enriched by the potential targets of the 10 most up-regulated (A) and 10 most down-regulated miRNAs (B). Expected and observed target counts and the FDR-corrected *P*-values are shown.

et al., 2006; Hong et al., 2010). Surprisingly, other members of this cluster rarely appear among dysregulated miRNAs, indicating that the primary transcript produces individual miRNAs that are subject to differential expression regulation via mechanisms affecting either processing or stability of the individual miRNAs. miR-210 has been functionally characterized as an important factor in regulation of the hypoxic response in tumorigenesis (Huang et al., 2009).

miR-101 was reported to be down-regulated and miR-21 up-regulated in five different studies. Expression of miR-101 was shown to be down-regulated in different cancer types and has been shown to possess tumor suppressor activity (Friedman et al., 2009; Wang et al., 2010). miR-21 is one of the most-studied oncogenic miRNAs and is implicated in modulation of the K-Ras pathway (Hatley et al., 2010). The expression levels of miR-21 in our study were very high in

TABLE 4. MiRNAs that Show Statistically Significant Altered Expression in NSCLC Tumor Tissue Versus Normal Lung in at Least Four Published Studies

Reference	Method	miRNAs <sup>a</sup>	miRNA									
			mir-9	mir-17	mir-21	mir-30a	mir-30d	mir-101	mir-106a	mir-126	miR-210	mir-451
This study	Microarray	858	↑	↑		↓		↓		↓	↑	↓
Gao et al., 2010	Microarray	555			↑	↓	↓	↓		↓		↓
Yang et al., 2010	Microarray	711				↓	↓	↓		↓		
Crawford et al., 2009	qRT-PCR	500	↑		↑	↓		↓		↓	↑	↓
Seike et al., 2009	Microarray	389			↑	↓	↓			↓	↑	↓
Raponi et al., 2009	Microarray	328		↑						↑	↑	
Navarro et al., 2009	qRT-PCR	18		↑						↑		
Yanaihara et al., 2006	Microarray	352	↓	↑	↑	↓		↓		↓	↑	
Volinia et al., 2006	Microarray	228	↑	↑	↑	↓	↓			↓	↑	
Hayashita et al., 2005	Northern, qRT-PCR	21		↑						↑		

<sup>a</sup>miRNAs, the number of different miRNAs investigated in the specific study; arrows indicate direction of change in expression.

Publications investigating only a single miRNA are omitted from this list. Prognostic markers for survival were analyzed by Gao et al., Raponi et al., and Yanaihara et al.

both normal and tumor samples. It could be that dysregulation of some highly expressed miRNAs (let-7 family or miR-21) escapes detection due to signal saturation and dynamic range limitations of the microarray. The greatest fold-change and statistical significance in our study was observed for miRNAs with moderate or low expression, such as miR-9, miR-205, and miR-206 (Fig. 1B).

Significantly altered expression of four miRNAs, miR-9, miR-30d, miR-106a, and miR-451, are reported in four studies. miR-451 is a very recently discovered miRNA, which can be synthesized via an alternative Dicer-independent biogenesis pathway (Cheloufi et al., 2010; Cifuentes et al., 2010) and plays a crucial role in erythropoiesis (Patrick et al., 2010; Yu et al., 2010). These data further indicate that additional miRNA profiling studies with better coverage of total human miRNAs have a good potential to reveal novel, functionally important small RNAs. For example, our study also identified significant changes in miR-1285, which was recently reported to directly target the tumor suppressor, *p53* (Tian et al., 2010).

One aspect that complicates miRNA expression analysis is the heterogeneous length of mature miRNA sequences generated from the same locus (Chiang et al., 2010). We were able to examine such heterogeneity because Illumina microarrays used in this study include a small number of probes against different species of the specific miRNAs. A few such miRNAs, which differ by one or two nucleotides according to different miR-Base releases, were altered in our analysis. For example, miR-151:9.1 (the probe 5'-ACTAGACT-GAAGCTCCTTGA-3' that is designed according

to miRBase version 9.1, February 2007) ranks 10th according to *P*-value in the list of altered genes in our analysis. However, use of a probe designed according to the current miRBase release, miR-151-3p (5'-CTAGACTGAAGCTCCTTGAGG-3', miRBase Release 16, September 2010), that differs by a few terminal basepairs does not show statistically significant changes in miR-151 expression. A second example is miR-17-5p probes, which also only show statistically significant changes in expression when designed using the miRBase version 9.1 (Supporting Information Table 1).

Most cancers are not a result of a single mutation but arise as a result of a complex interaction between environment and perturbations in genetic networks (Quigley and Balmain, 2009). Tumorigenic miRNA:mRNA interactions are often part of these complex pathways, where a single miRNA may have many targets and a specific mRNA may be regulated by multiple different miRNAs. To understand the impact of miRNA dysregulation on cellular physiology, we have analyzed the pathways most likely affected by aberrantly regulated miRNAs using combined effect of the 10 most significantly altered miRNAs (Fig. 3). It appears that several pathways associated with cellular motility, differentiation, and proliferation are most significantly enriched within the potential targets of miRNAs regulated in NSCLC samples. MAPK signaling and axonal guidance pathways have been reported to be associated with miRNA dysregulation in lung cancer previously (Raponi et al., 2009).

Our analysis showed that low expression of miR-374a is associated with a prognosis of poor survival. This miRNA, located on chromosome Xq13.2, has

not been previously shown to be associated with patient survival and its function is currently unknown. Earlier studies that have analyzed miRNAs with potential prognostic value have identified largely nonoverlapping sets of miRNAs. Rigorous analysis by Yu et al. reported a five-miRNA signature that predicts cancer relapse and survival of patients with lung cancer (miR-221 and let-7a expression correlates with survival, and expression of miR-137, miR-372, and miR-182\* correlates with poor clinical outcome) (Yu et al., 2008). A more recent study of 440 miRNAs in 107 male smokers (Landi et al., 2010) found a different signature; decreased expression of miR-25, miR-34c-5p, miR-191, let-7e, and miR-34a was associated with significantly poorer survival in SCC patients. Two additional studies (Raponi et al., 2009; Yanaihara et al., 2006) have also identified several miRNAs with prognostic value in lung cancer, but only miR-191, miR-155, and let-7a appear in more than one of the four aforementioned studies. We were able to detect expression changes in miR-191, miR-155, and let-7a in our study, but there was no correlation with survival (Supporting Information Table 2). The inconsistencies between studies could be due to technical details (custom or commercial microarray vs. qRT-PCR), the number of miRNAs screened and/or clinicopathological characteristics of the patients. However, such low reproducibility between studies indicates that further development of miRNA detection platforms and analysis methods is still needed before they will become practical in clinical settings.

In conclusion, we have identified an miRNA expression profile in early stage NSCLC and confirmed the altered expression of several previously identified tumor-associated miRNAs, including miR-9, miR-205, and miR-210. We also identified miR-374a as a novel potential marker for predicting a patient's prognosis. The analysis for potential targets of these dysregulated miRNAs suggests that several of them may be directly involved in tumor development and progression via cell growth, adhesion, and signaling pathways. Detailed analysis and experimental verification of predicted miRNA:mRNA interactions within these pathways is currently underway.

#### ACKNOWLEDGMENTS

The authors thank Viljo Soo for excellent technical assistance and Drs. Ana Rebanc and Märt Möls for fruitful discussions.

#### REFERENCES

- Calin G, Liu C, Sevignani C, Ferracin M, Felli N, Dumitru C, Shimizu M, Cimmino A, Zupo S, Dono M, Dell'Aquila M, Alder H, Rassenti L, Kipps T, Bullrich F, Negrini M, Croce C. 2004. MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. *Proc Natl Acad Sci USA* 101:11755–11760.
- Cheloufi S, Dos Santos C, Chong M, Hannon G. 2010. A dicer-independent miRNA biogenesis pathway that requires Ago catalysis. *Nature* 465:584–589.
- Chen X, Gong J, Zeng H, Chen N, Huang R, Huang Y, Nie L, Xu M, Xia J, Zhao F, Meng W, Zhou Q. 2010. MicroRNA145 targets BNIP3 and suppresses prostate cancer progression. *Cancer Res* 70:2728–2738.
- Chiang H, Schoenfeld L, Ruby J, Auyeung V, Spies N, Baek D, Johnston W, Russ C, Luo S, Babiarz J, Blalock R, Schroth G, Nusbaum C, Bartel D. 2010. Mammalian microRNAs: experimental evaluation of novel and previously annotated genes. *Genes Dev* 24:992–1009.
- Cifuentes D, Xue H, Taylor D, Patnode H, Mishima Y, Cheloufi S, Ma E, Mane S, Hannon G, Lawson N, Wolfe S, Giraldez A. 2010. A novel miRNA processing pathway independent of Dicer requires Argonaute2 catalytic activity. *Science* 328:1694–1698.
- Crawford M, Batte K, Yu L, Wu X, Nuovo G, Marsh C, Otterson G, Nana-Sinkam S. 2009. MicroRNA 133B targets pro-survival molecules MCL-1 and BCL2L2 in lung cancer. *Biochem Biophys Res Commun* 388:483–489.
- Croce C. 2009. Causes and consequences of microRNA dysregulation in cancer. *Nat Rev Genet* 10:704–714.
- Davoren P, McNeill R, Lowery A, Kerin M, Miller N. 2008. Identification of suitable endogenous control genes for microRNA gene expression analysis in human breast cancer. *BMC Mol Biol* 9:76.
- Dews M, Homayouni A, Yu D, Murphy D, Sevignani C, Wentzel E, Furth E, Lee W, Enders G, Mendell J, Thomas-Tikhonenko A. 2006. Augmentation of tumor angiogenesis by a Myc-activated microRNA cluster. *Nat Genet* 38:1060–1065.
- Du P, Kibbe W, Lin S. 2008. Lumi: A pipeline for processing Illumina microarray. *Bioinformatics* 24:1547–1548.
- Ebbhardt H, Tsang H, Dai D, Liu Y, Bostan B, Fahlman R. 2009. Meta-analysis of small RNA-sequencing errors reveals ubiquitous post-transcriptional RNA modifications. *Nucleic Acids Res* 37:2461–2470.
- Fabian M, Sonenberg N, Filipowicz W. 2010. Regulation of mRNA translation and stability by microRNAs. *Annu Rev Biochem* 79:351–379.
- Fish J, Santoro M, Morton S, Yu S, Yeh R, Wythe J, Ivey K, Bruneau B, Stainier D, Srivastava D. 2008. miR-126 regulates angiogenic signaling and vascular integrity. *Dev Cell* 15:272–284.
- Friedman J, Liang G, Liu C, Wolff E, Tsai Y, Ye W, Zhou X, Jones P. 2009. The putative tumor suppressor microRNA-101 modulates the cancer epigenome by repressing the polycomb group protein EZH2. *Cancer Res* 69:2623–2629.
- Gao W, Yu Y, Cao H, Shen H, Li X, Pan S, Shu Y. 2010. Deregulated expression of miR-21, miR-143 and miR-181a in non small cell lung cancer is related to clinicopathologic characteristics or patient prognosis. *Biomed Pharmacother* 64:399–408.
- Hand N, Master Z, Eauclaire S, Weinblatt D, Matthews R, Friedman J. 2009. The microRNA-30 family is required for vertebrate hepatobiliary development. *Gastroenterology* 136:1081–1090.
- Hatley M, Patrick D, Garcia M, Richardson J, Bassel-Duby R, van Rooij E, Olson E. 2010. Modulation of K-Ras-dependent lung tumorigenesis by MicroRNA-21. *Cancer Cell* 18:282–293.
- Hayashita Y, Osada H, Tatematsu Y, Yamada H, Yanagisawa K, Tomida S, Yatabe Y, Kawahara K, Sekido Y, Takahashi T. 2005. A polycistronic microRNA cluster, miR17-92, is overexpressed in human lung cancers and enhances cell proliferation. *Cancer Res* 65:9628–9632.
- Hong L, Lai M, Chen M, Xie C, Liao R, Kang Y, Xiao C, Hu W, Han J, Sun P. 2010. The miR-17-92 cluster of microRNAs confers tumorigenicity by inhibiting oncogene-induced senescence. *Cancer Res* 70:8547–8557.
- Huang X, Ding L, Bennewith K, Tong R, Welford S, Ang K, Story M, Le Q, Giaccia A. 2009. Hypoxia-inducible mir-210 regulates normoxic gene expression involved in tumor initiation. *Mol Cell* 35:856–867.
- Jemal A, Siegel R, Xu J, Ward E. 2010. Cancer statistics, 2010. *CA Cancer J Clin* 60:277–300.

- Kuhnert F, Mancuso M, Hampton J, Stankunas K, Asano T, Chen C, Kuo C. 2008. Attribution of vascular phenotypes of the murine *Egfl7* locus to the microRNA miR-126. *Development* 135:3989–3993.
- Kumar M, Erkeland S, Pester R, Chen C, Ebert M, Sharp P, Jacks T. 2008. Suppression of non-small cell lung tumor development by the let-7 microRNA family. *Proc Natl Acad Sci USA* 105:3903–3908.
- Landi M, Zhao Y, Rotunno M, Koshiol J, Liu H, Bergen A, Rubagotti M, Goldstein A, Linnoila I, Marincola F, Tucker M, Bertazzi P, Pesatori A, Caporaso N, McShane L, Wang E. 2010. MicroRNA expression differentiates histology and predicts survival of lung cancer. *Clin Cancer Res* 16:430–441.
- Liu B, Peng X, Zheng X, Wang J, Qin Y. 2009. MiR-126 restoration down-regulate VEGF and inhibit the growth of lung cancer cell lines in vitro and in vivo. *Lung Cancer* 66:169–175.
- Ma L, Young J, Prabhala H, Pan E, Mestdagh P, Muth D, Teruya-Feldstein J, Reinhardt F, Onder T, Valastyan S, Westermann F, Speleman F, Vandesompele J, Weinberg R. 2010. miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis. *Nat Cell Biol* 12:247–256.
- Miska E, Alvarez-Saavedra E, Townsend M, Yoshii A, Sestan N, Rakic P, Constantine-Paton M, Horvitz H. 2004. Microarray analysis of microRNA expression in the developing mammalian brain. *Genome Biol* 5:R68.
- Mitchell P, Parkin R, Kroh E, Fritz B, Wyman S, Pogosova-Agadjanyan E, Peterson A, Noteboom J, O'Briant K, Allen A, Lin D, Urban N, Drescher C, Knudsen B, Stirewalt D, Gentleman R, Vessella R, Nelson P, Martin D, Tewari M. 2008. Circulating microRNAs as stable blood-based markers for cancer detection. *Proc Natl Acad Sci USA* 105:10513–10518.
- Mountain CF. 2000. The international system for staging lung cancer. *Semin Surg Oncol* 18:106–115.
- Navarro A, Marrades R, Viñolas N, Quera A, Agustí C, Huerta A, Ramirez J, Torres A, Monzo M. 2009. MicroRNAs expressed during lung cancer development are expressed in human pseudoglandular lung embryogenesis. *Oncology* 76:162–169.
- Nelis M, Esko T, Mägi R, Zimprich F, Zimprich A, Toncheva D, Karachanak S, Piskáčeková T, Balascák I, Peltonen L, Jakkula E, Rehnström K, Lathrop M, Heath S, Galan P, Schreiber S, Meitinger T, Pfeufer A, Wichmann H, Melegh B, Polgár N, Toniolo D, Gasparini P, D'Adamo P, Klovins J, Nikitina-Zake L, Kucinskis V, Kasnauskienė J, Lubinski J, Debnjak T, Limborska S, Khrunin A, Estivill X, Rabionet R, Marsal S, Julià A, Antonarakis S, Deutsch S, Borel C, Attar H, Gagnebin M, Macek M, Krawczak M, Remm M, Metspalu A. 2009. Genetic structure of Europeans: A view from the North-East. *PLoS One* 4:e5472.
- Patrick D, Zhang C, Tao Y, Yao H, Qi X, Schwartz R, Jun-Shen Huang L, Olson E. 2010. Defective erythroid differentiation in miR-451 mutant mice mediated by 14-3-3zeta. *Genes Dev* 24:1614–1619.
- Pineau P, Volinia S, McJunkin K, Marchio A, Battiston C, Terris B, Mazzaferro V, Lowe S, Croce C, Dejean A. 2010. miR-221 overexpression contributes to liver tumorigenesis. *Proc Natl Acad Sci USA* 107:264–269.
- Quigley D, Balmain A. 2009. Systems genetics analysis of cancer susceptibility: From mouse models to humans. *Nat Rev Genet* 10:651–657.
- Raponi M, Dossey L, Jatkoec T, Wu X, Chen G, Fan H, Beer D. 2009. MicroRNA classifiers for predicting prognosis of squamous cell lung cancer. *Cancer Res* 69:5776–5783.
- Reinhart B, Slack F, Basson M, Pasquinelli A, Bettinger J, Rougvie A, Horvitz H, Ruvkun G. 2000. The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature* 403:901–906.
- Ryan B, Robles A, Harris C. 2010. Genetic variation in microRNA networks: the implications for cancer research. *Nat Rev Cancer* 10:389–402.
- Schaefer A, Stephan C, Busch J, Yousef G, Jung K. 2010. Diagnostic, prognostic and therapeutic implications of microRNAs in urologic tumors. *Nat Rev Urol* 7:286–297.
- Seike M, Goto A, Okano T, Bowman E, Schetter A, Horikawa I, Mathe E, Jen J, Yang P, Sugimura H, Gemma A, Kudoh S, Croce C, Harris C. 2009. MiR-21 is an EGF-regulated anti-apoptotic factor in lung cancer in never-smokers. *Proc Natl Acad Sci USA* 106:12085–12090.
- Song G, Zhang Y, Wang L. 2009. MicroRNA-206 targets notch3, activates apoptosis, and inhibits tumor cell migration and focus formation. *J Biol Chem* 284:31921–31927.
- Takamizawa J, Konishi H, Yanagisawa K, Tomida S, Osada H, Endoh H, Harano T, Yatabe Y, Nagino M, Nimura Y, Mitsudomi T, Takahashi T. 2004. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. *Cancer Res* 64:3753–3756.
- Thomson J, Parker J, Perou C, Hammond S. 2004. A custom microarray platform for analysis of microRNA gene expression. *Nat Methods* 1:47–53.
- Tian S, Huang S, Wu S, Guo W, Li J, He X. 2010. MicroRNA-1285 inhibits the expression of p53 by directly targeting its 3' untranslated region. *Biochem Biophys Res Commun* 396:435–439.
- Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 3:RESEARCH0034.
- Valk K, Vooder T, Kolde R, Reintam MA, Petzold C, Vilo J, Metspalu A. 2010. Gene expression profiles of non-small cell lung cancer: survival prediction and new biomarkers. *Oncology* 79:283–292.
- Volinia S, Calin G, Liu C, Ambs S, Cimmino A, Petrocca F, Visone R, Iorio M, Roldo C, Ferracin M, Prueitt R, Yanaihara N, Lanza G, Scarpa A, Vecchione A, Negrini M, Harris C, Croce C. 2006. A microRNA expression signature of human solid tumors defines cancer gene targets. *Proc Natl Acad Sci USA* 103:2257–2261.
- Wang H, Ruan H, He X, Ma Y, Jiang X, Xia Y, Ye Z, Tao H. 2010. MicroRNA-101 is down-regulated in gastric cancer and involved in cell migration and invasion. *Eur J Cancer* 46:2295–2303.
- Yanaihara N, Caplen N, Bowman E, Seike M, Kumamoto K, Yi M, Stephens R, Okamoto A, Yokota J, Tanaka T, Calin G, Liu C, Croce C, Harris C. 2006. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. *Cancer Cell* 9:189–198.
- Yang Y, Li X, Yang Q, Wang X, Zhou Y, Jiang T, Ma Q, Wang Y. 2010. The role of microRNA in human lung squamous cell carcinoma. *Cancer Genet Cytogenet* 200:127–133.
- Yu D, dos Santos C, Zhao G, Jiang J, Amigo J, Khandros E, Dore L, Yao Y, D'Souza J, Zhang Z, Ghaffari S, Choi J, Friend S, Tong W, Orange J, Paw B, Weiss M. 2010. miR-451 protects against erythroid oxidant stress by repressing 14-3-3zeta. *Genes Dev* 24:1620–1633.
- Yu S, Chen H, Chang G, Chen C, Chen H, Singh S, Cheng C, Yu C, Lee Y, Chen H, Su T, Chiang C, Li H, Hong Q, Su H, Chen C, Chen W, Liu C, Chan W, Li K, Chen J, Yang P. 2008. MicroRNA signature predicts survival and relapse in lung cancer. *Cancer Cell* 13:48–57.