Unique microRNA molecular profiles in lung cancer diagnosis and prognosis

Nozomu Yanaihara,[1] Natasha Caplen,[2] Elise Bowman,[1] Masahiro Seike,[1] Kensuke Kumamoto,[1] Ming Yi,[3] Robert M. Stephens,[3] Aikou Okamoto,[4] Jun Yokota,[5] Tadao Tanaka,[4] George Adrian Calin,[6] Chang-Gong Liu,[6] Carlo M. Croce,[6] and Curtis C. Harris[1,*]

[1] Laboratory of Human Carcinogenesis, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892
[2] Gene Silencing Section, Office of Science and Technology Partnership, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20892
[3] Advanced Biomedical Computing Center, National Cancer Institute-Frederick/SAIC-Frederick Inc., Frederick, Maryland 21702
[4] Department of Obstetrics and Gynecology, The Jikei University School of Medicine, Tokyo 105-8461, Japan
[5] Biology Division, National Cancer Center Research Institute, Tokyo 104-0045, Japan
[6] Molecular Virology, Immunology and Medical Genetics, The Ohio State University Comprehensive Cancer Center, Columbus, Ohio 43210
*Correspondence: curtis_harris@nih.gov

Summary

MicroRNA (miRNA) expression profiles for lung cancers were examined to investigate miRNA's involvement in lung carcinogenesis. miRNA microarray analysis identified statistical unique profiles, which could discriminate lung cancers from noncancerous lung tissues as well as molecular signatures that differ in tumor histology. miRNA expression profiles correlated with survival of lung adenocarcinomas, including those classified as disease stage I. High *hsa-mir-155* and low *hsa-let-7a-2* expression correlated with poor survival by univariate analysis as well as multivariate analysis for *hsa-mir-155*. The miRNA expression signature on outcome was confirmed by real-time RT-PCR analysis of precursor miRNAs and crossvalidated with an independent set of adenocarcinomas. These results indicate that miRNA expression profiles are diagnostic and prognostic markers of lung cancer.

Introduction

Lung cancer is the leading cause of cancer deaths in the world, and its etiology is primarily genetic and epigenetic damage caused by tobacco smoke (Travis et al., 2004). Systematic analysis of mRNA and protein expression levels among thousands of genes has also contributed to defining the molecular network of lung carcinogenesis (Meyerson and Carbone, 2005; Granville and Dennis, 2005). Defects in both the *p53* and *RB*/*p16* pathways are common in lung cancer. Several other genes, such as *K-ras*, *PTEN*, *FHIT*, and *MYO18B*, are genetically altered, though less frequently (Minna et al., 2002; Sekido et al., 2003; Yokota and Kohno, 2004). Although focusing on known genes and proteins has already yielded new information, previously unknown markers such as noncoding RNA gene products may also lend insight into the biology of lung cancer.

MicroRNAs (miRNAs) are small noncoding RNA gene products about 22 nt long that are found in diverse organisms and play key roles in regulating the translation and degradation of mRNAs through base pairing to partially complementary sites, predominately in the untranslated region of the message (Lagos-Quintana et al., 2001; Lau et al., 2001; Lee and Ambros, 2001). miRNAs are expressed as long precursor RNAs that are processed by a cellular nuclease, Drosha, before being transported by an Exportin-5-dependent mechanism into the cytoplasm (Yi et al., 2003; Gregory and Shiekhattar, 2005). Once in the cytoplasm, miRNAs are cleaved further by the enzyme DICER (Lee et al., 2002, 2003), and this results in 17–24 nt miRNAs that are associated with a cellular complex that is similar to the RNA-induced silencing complex that participates in RNA interference (Hutvagner and Zamore, 2002). Recently, it was reported that *DICER* expression levels were reduced in a fraction of lung

SIGNIFICANCE

miRNAs are a class of small noncoding RNA genes found to be abnormally expressed in several types of cancer, suggesting that miRNAs play a substantial role in the pathogenesis of human cancers. Lung cancer is the leading cause of cancer deaths in the world, reflecting the need for a better understanding of the mechanisms that underlie lung carcinogenesis. Although focusing on known genes and proteins has already yielded new information, unknown markers may also lend insight into the biology of lung cancer. We showed that lung cancer has extensive alterations of miRNA expression that may deregulate cancer-related genes. Furthermore, the miRNA molecular signature of lung adenocarcinomas, including those without evidence of metastasis, also correlates with patient survival.

cancers with a significant prognostic impact on patient survival (Karube et al., 2005). The biological functions of most miRNAs are not yet fully understood. It has been suggested that the miRNAs are involved in various biological processes, including cell proliferation, cell death, stress resistance, and fat metabolism, through the regulation of gene expression (Ambros, 2003).

Our understanding of miRNA expression patterns and function in normal or neoplastic human cells is just starting to emerge. miRNA genes are frequently located at fragile sites (FRAs), as well as in minimal regions of loss of heterozygosity, minimal regions of amplification, or common breakpoint regions, suggesting that miRNAs might be a new class of genes involved in human tumorigenesis (Calin et al., 2004b). For example, *mir-15-a* and *mir-16-1* are frequently deleted and/or downregulated in patients with B cell chronic lymphocytic leukemia (Calin et al., 2002). Other links between cancer and miRNA have been reported, including reduced expression of *mir-143* and *mir-145* in colorectal cancers (Michael et al., 2003) and *let-7* in lung cancers (Takamizawa et al., 2004), high expression of the precursor *mir-155* in Burkitt's lymphomas (Metzler et al., 2004), and oncogenic function of *mir-17-92* cluster in human B cell lymphomas as well as in lung cancers (He et al., 2005; Hayashita et al., 2005). The precise mechanisms regulating miRNA expression are unknown. However, several mechanisms, including genetic and epigenetic alteration, might affect miRNA expression, and they may lead to alterations in the pattern of target genes expression in cancers. It was shown that miRNA expression patterns have relevance to the biological and clinical behavior of human B cell chronic lymphocytic leukemia and solid tumors, including breast cancers (Calin et al., 2004a; Iorio et al., 2005; Volinia et al., 2006). One or more members of the *let-7* family regulate RAS expression in human cells, and thus, *let-7* may play a major role in human lung carcinogenesis as a tumor suppressor gene (Johnson et al., 2005). Recently, miRNA expression profiles have been shown to be potential tools for cancer diagnosis (Lu et al., 2005). These and other data are consistent with the hypothesis that miRNAs play a substantial role in the pathogenesis of human cancers.

In this study, we investigated the miRNA expression profiles in human lung cancer and miRNA regulation by epigenetic mechanisms and found that the miRNA molecular profile of lung adenocarcinoma correlates with patient survival.

Results

Altered miRNA expression in primary lung cancers and identification of miRNAs associated with clinicopathological features of lung cancer

We analyzed the miRNA expression in 104 pairs of primary lung cancers and corresponding noncancerous lung tissues. We compared miRNA expression of several group pairs as listed in Table 1. Expression profiles were generated by comparing lung cancers, except when comparing lung cancer tissues with corresponding noncancerous lung tissues. We identified miRNAs, which were expressed differently in phenotypical and histological classifications (Table 1). When we compared miRNA expression among lung cancer tissues versus corresponding noncancerous lung tissues, 43 miRNAs had statistical differences in expression between groups (Table 2). In class comparison analysis using our microarray analysis tool, the multivariate permutation test was performed to control multiple comparisons. It provides a specific confidence level for ensuring that the number of false discoveries does not exceed a target level or for ensuring that the proportion of the gene list that are false discoveries does not exceed a target level. Thus, the probability of identifying at least 43 miRNAs by chance at the <0.001 level, if there are no real differences between the classes, was 0 as estimated by the multivariate permutation test. Furthermore, 91% of 104 lung cancers were correctly classified using the leave-one-out crossvalidated class prediction method based on the compound covariate predictor. Based on 2000 random permutations, the p value, which is defined as the proportion of the random permutations that gave a crossvalidated error rate no greater than the crossvalidated error rate with the real data, was < 0.0005.

Several of these miRNAs were associated with FRAs (Table 2). In particular, three miRNAs are located inside FRAs (*hsa-mir-21* at FRA17B, *hsa-mir-27b* at FRA9D, and *hsa-mir-32* at FRA9E). Furthermore, many of these miRNAs are located at frequently deleted or amplified regions in several malignancies (Table 2). For example, *hsa-mir-21* and *hsa-mir-205* are located at the region amplified in lung cancer, whereas *hsa-mir-126** and *hsa-mir-126* are at 9q34.3, a region deleted in lung cancer. Reduced expression of precursor *let-7a-2* and *let-7f-1* was also found in adenocarcinoma and squamous cell carcinoma at a p value cutoff of 0.05, respectively. In the same way, comparison analyses between lung adenocarcinoma versus noncancerous tissues and squamous cell carcinoma versus noncancerous tissues revealed 17 and 16 miRNAs with statistically different expression, respectively (Table S2 in the Supplemental Data available with this article online). Six miRNAs (*hsa-mir-21*, *hsa-mir-191*, *hsa-mir-155*, *hsa-mir-210*, *has-mir-126**, and *hsa-mir-224*) were shared in both histological types of non-small cell lung carcinoma (NSCLC).

Next, we asked whether the microarray data revealed specific molecular signatures for subsets of lung cancer that differ in clinical behavior. For this analysis, we examined the relationship of five types of clinical and pathological information (Table 1). Among them, we identified six miRNAs (*hsa-mir-205*, *hsa-mir-99b*, *hsa-mir-203*, *hsa-mir-202*, *hsa-mir-102*, and *hsa-mir-204-prec*) that were expressed differently in the two most common histological types of NSCLC, adenocarcinoma and squamous cell carcinoma. The expression levels of *hsa-mir-99b* and *hsa-mir-102* were higher in adenocarcinoma. No miRNAs were identified as differently expressed when classified by age, gender, or race in our data set.

Validation of the microarray data by the solution hybridization detection method and real-time RT-PCR analysis

We used the solution hybridization detection method for mature miRNAs and real-time RT-PCR analysis for precursor miRNAs to validate the results from microarray analysis. First, the microarray data of three miRNAs (*hsa-mir-21*, *hsa-mir-126**, and *hsa-mir-205*) were analyzed by the solution hybridization detection method. Seven pairs of primary lung cancers and corresponding noncancerous lung tissues in which sufficient RNA was available were analyzed. The mature forms of *hsa-mir-21* and *hsa-mir-205* were clearly upregulated in lung cancer tissues when compared with the corresponding noncancerous lung tissues (Figure 1A). In contrast, *hsa-mir-126** was downregulated in most of the lung cancer tissues examined. Therefore, the analyses confirmed the microarray data for these three miRNAs.

Table 1. Comparison analysis of clinicopathological classifications

Classification (number)	Total	No. of genes	FDR	Percent correctly classified (p value)
Phenotypical classification				
All tumor (104) versus all normal (104)	208	43	0	91 (<0.0005)
Adeno tumor (65) versus adeno normal (65)	130	17	0.001	80 (<0.0005)
SCC tumor (39) versus SCC normal (39)	78	16	0	92 (<0.0005)
Histological classification				
Adeno tumor (65) versus SCC tumor (39)	104	6	0.001	81 (<0.0005)
Age classification				
All; age < 67 (56) versus age ≥ 67 (48)	104	0		
Adeno; age < 67 (37) versus age ≥ 67 (28)	65	0		
SCC; age < 67 (19) versus age ≥ 67 (20)	39	0		
Sex classification				
All; male (65) versus female (39)	104	0		
Adeno; male (39) versus female (26)	65	0		
SCC; male (26) versus female (13)	39	0		
Race classification				
All; African American (21) versus Caucasian (83)	104	0		
Adeno; African American (13) versus Caucasian (52)	65	0		
SCC; African American (8) versus Caucasian (31)	39	0		
Stage classification				
All; stage I (65) versus stage II (17) versus stages III and IV (22)	104	0		
Adeno; stage I (41) versus stage II (8) versus stages III and IV (16)	65	1		
SCC; stage I (24) versus stage II (9) versus stages III and IV (6)	39	0		

No. of genes, number of genes significant at 0.001; FDR, false discovery rate, which is probability of significant genes by chance. Percent correctly classified (p value), the leave-one-out crossvalidated class prediction method based on the compound covariate predictor. The p value is the proportion of the random permutations that gave a crossvalidated error rate no greater than the crossvalidated error rate with the real data. Adeno, adenocarcinoma; SCC, squamous cell carcinoma.

We also performed real-time RT-PCR analysis of precursor miRNAs. First, cDNA from 16 pairs of lung adenocarcinoma and 16 pairs of lung squamous cell carcinoma was prepared by gene-specific primers (*hsa-mir-21*, *hsa-mir-126**, *hsa-mir-205*, and *U6*), and then, real-time RT-PCR analysis for these miRNAs and an endogenous control were performed. At least 2-fold upregulation of precursor *hsa-mir-21* and *hsa-mir-205* expression was found in 66% and 56% out of 32 cases, respectively, when compared with that in the corresponding noncancerous tissues. The differences were statistically significant at $p < 0.001$ by paired Student's t test (Figure 1B). On the other hand, 31% of 32 cases examined were found to exhibit >50% reduction in precursor *hsa-mir-126** expression even though the reduction was not statistically significant (Figure 1B). These findings show the frequent occurrence of either upregulation or a reduction of specific precursor miRNAs in lung cancers, as was seen in the mature miRNAs by using microarray analysis.

Correlation between miRNA expression profiles and prognosis of lung adenocarcinoma patients

We next investigated the correlation of miRNA expression profiles with patient survival. A univariate Cox proportional hazard regression model with global permutation test in BRB-ArrayTools indicated that eight miRNAs (*hsa-mir-155*, *hsa-mir-17-3p*, *hsa-mir-106a*, *hsa-mir-93*, *hsa-let-7a-2*, *hsa-mir-145*, *hsa-let-7b*, and *hsa-mir-21*) were related to the adenocarcinoma patient's survival. Patients with high expression of either *hsa-mir-155*, *hsa-mir-17-3p*, *hsa-mir-106a*, *hsa-mir-93*, or *hsa-mir-21* and low expression of either *hsa-let-7a-2*, *hsa-let-7b*, or *hsa-mir-145* were found to have a significantly worse prognosis. In addition, the survival analysis among the 41 stage I adenocarcinoma patients revealed that three miRNAs (*hsa-mir-155*, *hsa-mir-17-3p*, and *hsa-mir-20*) were associated with patient outcome. This indicated the important relationship between miRNA expression profiles and patient survival, independent of disease stage.

Because five miRNAs (*hsa-mir-155*, *hsa-mir-17-3p*, *hsa-let-7a-2*, *hsa-mir-145*, and *hsa-mir-21*) out of these miRNAs were expressed differently among lung cancer tissues versus corresponding noncancerous lung tissues, these miRNAs were used for further survival analysis. The ratio of lung cancer expression to corresponding noncancerous lung tissue expression for each of these five miRNAs was calculated, and the cases were classified according to the expression ratio. Using these groupings for each miRNA, Kaplan-Meier survival analysis was performed. Kaplan-Meier survival estimates showed that the lung adenocarcinoma patients with either high *hsa-mir-155* or reduced *hsa-let-7a-2* expression had a poorer survival than the patients with low *hsa-mir-155* or high *hsa-let-7a-2* expression, respectively (Figure 2). The difference in the prognosis of these two groups was statistically significant for *hsa-mir-155* ($p = 0.006$; log-rank test) and marginally significant for *hsa-let-7a-2* ($p = 0.033$; log-rank test). Survival analysis of the clinicopathological factors showed that stage was associated with survival ($p = 0.01$; log-rank test), while age, sex, and smoking history did not account for poor prognosis (Table 3). To adjust for multiple comparisons, we used the method by Storey and Tibshirani (2003), limiting the false discovery rates to 0.05. When this rate was used, *hsa-mir-155* and disease stage were still statistically significant. Subsequently, a multivariate Cox

Table 2. Forty-three miRNAs differentially expressed in lung cancer tissues versus noncancerous lung tissues

miRNA	Location	p value	Type	FRA association[a]	Cancer-associated genomic regions[a]	Host gene[b]
hsa-mir-21	17q23.2	<1e-07	up	FRA17B	amp-neuroblastoma; lung cancer	*TMEM49*
hsa-mir-191	3p21.31	<1e-07	up			unknown
*hsa-mir-126**	9q34.3	<1e-07	down		del-NSCLC; HCC	*EGFL-7*
hsa-mir-210	11p15.5	1.00e-07	up		del-ovarian; lung cancer	unknown
hsa-mir-155	21q21.3	1.00e-07	up		amp-colon cancer	*BIC*
hsa-mir-143	5q32	4.00e-07	down		del-prostate cancer	mIncRNA
hsa-mir-205	1q32.2	4.00e-07	up		amp-lung cancer	mIncRNA
hsa-mir-192-prec	11q13.1	5.00e-07	down	FRA11A	del-thyroid cancer	mIncRNA
hsa-mir-224	Xq28	5.00e-07	down	FRAXF		*GABRE*
hsa-mir-126	9q34.3	7.00e-07	down		del-NSCLC; HCC	*EGFL-7*
hsa-mir-24-2	19p13.1	1.30e-06	up			ND
hsa-mir-30a-5p	6q13	4.80e-06	down			mIncRNA
hsa-mir-212	17p13.3	5.00e-06	up			ND
hsa-mir-140	16q22.1	5.10e-06	down			*ATROPIN-1*
hsa-mir-9	15q26.1	6.50e-06	down			unknown
hsa-mir-214	1q24.3	8.60e-06	up			ND
hsa-mir-17-3p	13q31.3	9.40e-06	up			unknown
hsa-mir-124a-1	8p23.1	1.23e-05	down		amp-MFHs	unknown
hsa-mir-218-2	5q34	1.34e-05	down			*SLIT3*
hsa-mir-95	4p16.1	1.48e-05	down			*ABLIM2*
hsa-mir-145	5q32	1.90e-05	down		del-prostate cancer	mIncRNA
hsa-mir-198	3q13.33	2.43e-05	down			*FSTL1*
hsa-mir-216-prec	2p16.1	3.05e-05	down			ND
hsa-mir-219-1	6p21.32	5.56e-05	down			ND
hsa-mir-106a	Xq26.2	6.20e-05	up		del-ovarian cancer	ND
hsa-mir-197	1p13.3	7.23e-05	up			ND
hsa-mir-192	11q13.1	0.000119	up	FRA11A	del-thyroid cancer	ND
hsa-mir-125a-prec	19q13.41	0.000143	down			mIncRNA
hsa-mir-26a-1-prec	3p22.3	0.000148	down		del-epithelial cancer	*NIF1*
hsa-mir-146	5q33.3	0.000163	up			mIncRNA
hsa-mir-203	14q32.33	0.000267	up			ND
hsa-mir-199b-prec	9q34.11	0.000304	down		del-bladder cancer	*GOLGA2*
hsa-let-7a-2-prec	11q24.1	0.000398	down	FRA11B	del-lung cancer	mIncRNA
hsa-mir-27b	9q22.32	0.000454	down	FRA9D	del-bladder cancer	unknown
hsa-mir-32	9q31.3	0.000458	down	FRA9E	del-lung cancer	unknown
hsa-mir-29b-2	1q32.2	0.000466	down			mIncRNA
hsa-mir-220	Xq25	0.000630	down			ND
hsa-mir-33	22q13.2	0.000683	down		del-colon cancer	*SREBF2*
hsa-mir-181c-prec	19p13.12	0.000736	down			*NANOS3*
hsa-mir-150	19q13.33	0.000784	up			ND
hsa-mir-101-1	1p31.3	0.000844	down	FRA1C	del-ovarian; breast cancer	ND
hsa-mir-124a-3	20q13.33	0.000968	down			ND
hsa-mir-125a	19q13.41	0.000993	down			ND

Amp, amplification; del, deletion; NSCLC, non-small cell lung carcinoma; HCC, hepatocellular carcinoma; mIncRNA, mRNA-like noncoding RNA; ND, not defined; MFHs, malignant fibrous histocytomas.
[a]Information was obtained from previous report (Calin et al., 2004b).
[b]Information was obtained from previous report (Rodriguez et al., 2004).

proportional hazard regression analysis using all of these clinicopathological and molecular factors indicated that high *hsa-mir-155* expression was a significantly unfavorable prognostic factor independent of other clinicopathological factors (p = 0.027; risk ratio 3.03; 95% confidence interval [CI], 1.13–8.14) in addition to disease stage (p = 0.013; risk ratio 3.27; 95% CI, 1.31–8.37) (Table 3).

To investigate the biological consequence of altered *hsa-mir-155* and *hsa-let-7a-2* expression, we conducted a bioinformatic analysis grouping the predicted targets of these miRNAs by Gene Ontology (GO) terms (Table S2). In addition to associations with more general functional GO terms, a significant enrichment for targets associated with transcription was seen for *hsa-mir-155*. *hsa-let-7a* showed an overrepresentation of gene targets linked with protein kinase and intracellular signaling cascades, a finding consistent with the reported functional interaction between *let-7* and RAS (Johnson et al., 2005).

Validation of miRNA expression signature on lung adenocarcinoma patient survival using an independent set of adenocarcinoma patients

Real-time RT-PCR analysis was performed for *hsa-mir-155* and *hsa-let-7a-2* to determine whether the precursor miRNA expression also had a prognostic impact on adenocarcinoma patients. First, 32 pairs of adenocarcinomas from the original set, in which RNA was available, were subjected to real-time RT-PCR analysis. The ratio of lung cancer expression to corresponding noncancerous lung tissue expression was calculated, and the cases were classified according to the expression ratio. Kaplan-Meier survival analysis (Figures S1A and S1B) demonstrated a significantly worse survival for patients with either high precursor *hsa-mir-155* expression (p = 0.047; log-rank test) or reduced precursor *hsa-let-7a-2* expression (p = 0.037; log-rank test) (Table 4). To further validate the prognosis classifiers described here, we analyzed an additional independent set

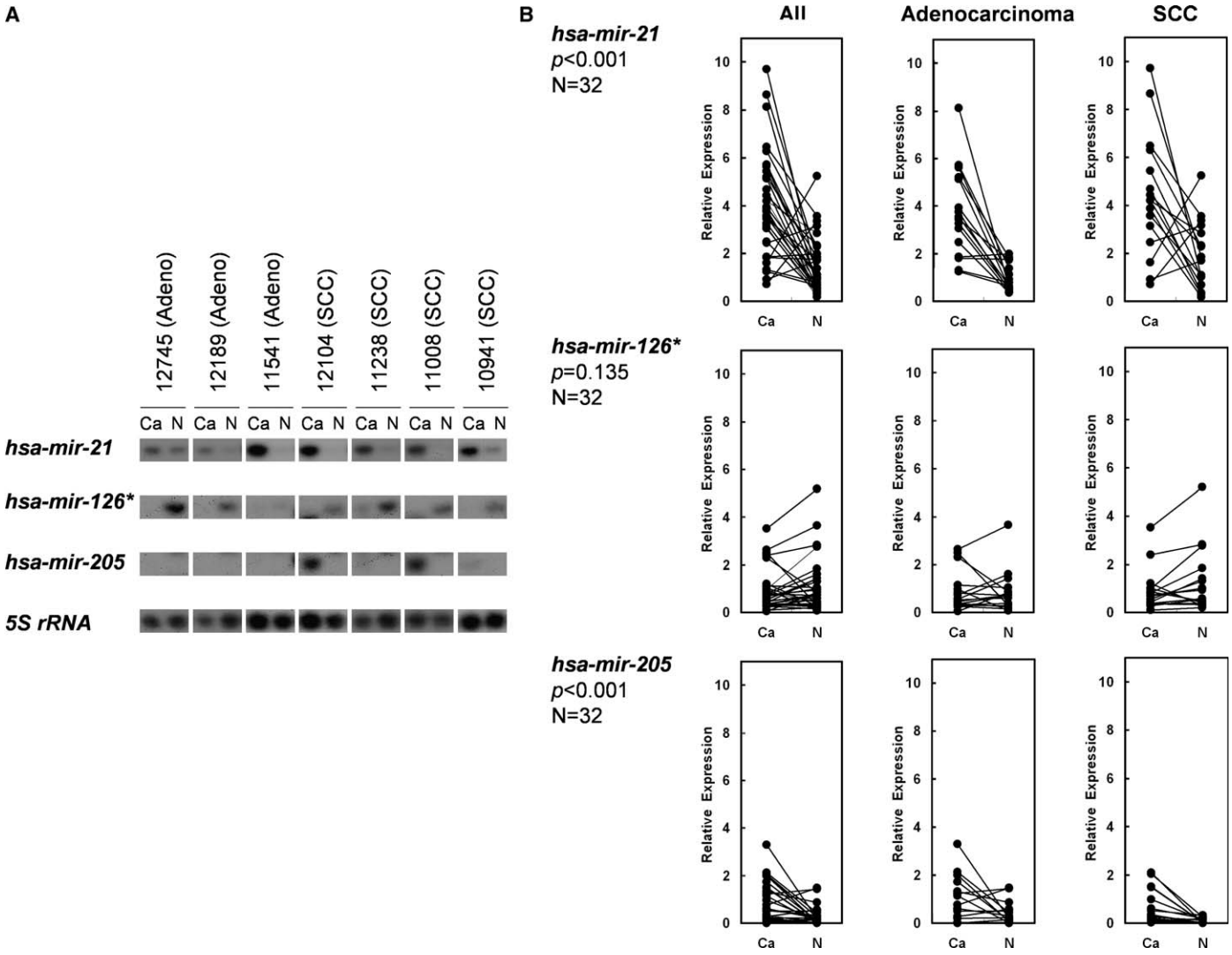

Figure 1. Validation of the microarray data by the solution hybridization detection method and real-time RT-PCR analysis

A: Solution hybridization detection of mature miRNA expression in lung cancers. Ca, cancerous tissues; N, noncancerous lung tissues. rRNA (5S) served as a loading control.

B: Real-time RT-PCR analysis of precursor miRNA expression in lung cancers. Ca, cancerous tissues; N, noncancerous lung tissues prepared from the same patient. Paired Student's t test was performed to ascertain statistical significance between the expression levels in cancer tissue and noncancerous tissues.

of 32 adenocarcinomas using real-time RT-PCR analysis. Kaplan-Meier survival curves (Figures S1C and S1D) showed a clear relationship in precursor *hsa-mir-155* expression (p = 0.033; log-rank test) and approached significance in *hsa-let-7a-2* expression (p = 0.084; log-rank test) in this cohort as well (Table 4). In addition, high precursor *hsa-mir-155* expression was found to be an independent predictor of poor prognosis by a multivariate Cox proportional hazard regression analysis (Table 4). To further confirm if there was any grouping bias between the original set and the additional set, we performed univariate and multivariate survival analyses for all 64 cases. Consistently, these analyses showed the significance of precursor *hsa-mir-155* expression (Table 4; Figure 3A). Of note, reduced precursor *hsa-let-7a-2* expression also had a similar prognostic impact on adenocarcinoma patients (Table 4; Figure 3B) consistent with a previous report (Takamizawa et al., 2004).

Lack of epigenetic regulation of miRNA expression in NSCLC cell lines

We employed miRNA microarray to analyze the changes in miRNA expression after 5-aza-2′-deoxycytidine (5-aza-dC) and/or Trichostatin A (TSA) treatment in two lung cancer cell lines, A549 and NCI-H157. Although we could confirm reexpression of a known transcriptional-silenced gene (*MYO18B*) as a positive control (Figure S2), none of the miRNAs showed a statistically significant change in increased expression after treatment with 5-aza-dC and/or TSA, suggesting that hypermethylation and histone deacetylation were not responsible for the reduced miRNA expression in at least these two cells.

Discussion

Little is known about the expression levels or function of miRNAs in normal and neoplastic cells, although it is becoming clear that

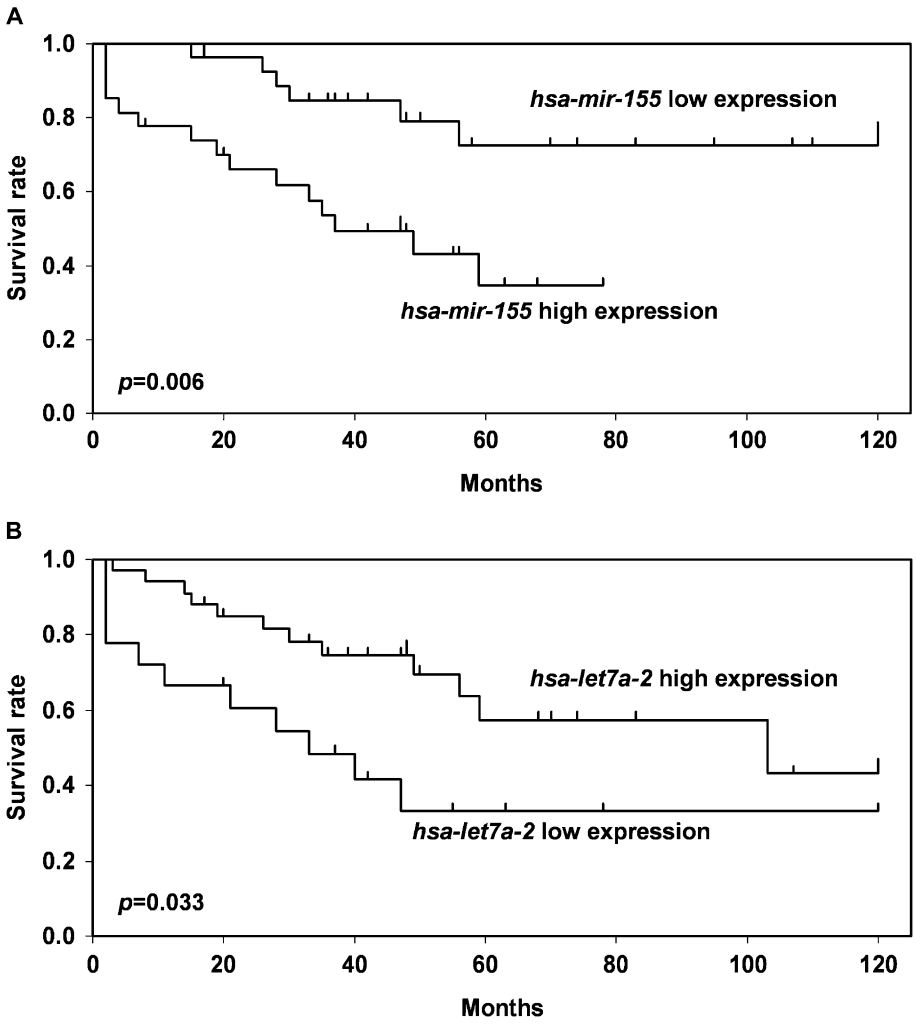

Figure 2. Kaplan-Meier survival curves for adenocarcinoma patients

Adenocarcinoma cases in which hybridization intensity was different from background (Experimental Procedures) were classified according to either *hsa-mir-155* expression or *hsa-let-7a-2* expression by miRNA microarray analysis. The survival data were compared using the log-rank test.

A: *hsa-mir-155* high expression (n = 27), group with the expression ratio ≥ mean expression ratio (1.42). The mean expression ratio is defined as mean expression ratio = mean of tumor expression/mean of noncancerous tissue expression. *hsa-mir-155* low expression (n = 28), group with the expression ratio < mean expression ratio.

B: *hsa-let-7a-2* high expression (n = 34), group with the expression ratio ≥ mean expression ratio (0.95). *hsa-let-7a-2* low expression (n = 18), group with the expression ratio < mean expression ratio.

miRNAs play major roles in the regulation of gene expression during development (Ambros, 2003; McManus, 2003). We reported here that the genome-wide expression profiling of miRNAs was significantly different among primary lung cancers and corresponding noncancerous lung tissues. The microarray data were validated by both solution hybridization detection method for mature miRNAs and real-time RT-PCR analysis for precursor miRNAs. Several of the miRNAs identified as differentially expressed are located inside FRAs and/or in the chromosomal regions where genomic imbalance in lung cancers has been observed previously with high frequency. As FRAs are preferential sites of translocation, deletion, amplification, or integration of exogenous genome, it is possible that miRNAs located near FRAs could be possible targets of such genomic alteration. Even though there is the possibility that the differences in miRNA expression may simply be a surrogate for cytogenetic changes in lung cancers, the fact that >50% of miRNAs are located at cancer-related chromosomal regions supported the idea that miRNAs may play a role as oncogenes or tumor suppressor genes. Moreover, these miRNAs are suggested to be involved in cancer. High expression of *mir-155* was found in Burkitt's lymphoma and B cell lymphomas (Metzler et al., 2004; Eis et al., 2005). It was also reported that *mir-143* and *mir-145* are reduced in colon cancer (Michael et al., 2003). The potential involvement of reduced *let-7* expression in lung cancers has been reported by two different groups (Takamizawa et al., 2004; Johnson et al., 2005). Consistent with a previous report, we also found reduced expression of precursor *hsa-let-7a-2* and *let-7f-1* in our data set. Overexpression of *mir-17-92* cluster was recently reported in lung cancers, especially in those with small cell lung cancers (Hayashita et al., 2005).

Precise molecular mechanisms for the altered expression of miRNAs in lung cancers are unclear. Abnormal expression of miRNAs in lung cancers could be caused by somatic genetic alterations. Alternatively, the reduced expression of miRNAs in lung cancer could be caused by epigenetic change such as DNA methylation and alterations of chromatin structure, which are important processes of transcriptional silencing in many genes, including tumor suppressor genes, and as an alternative to genetic defects in human carcinogenesis (Jones and Baylin, 2002; Eberharter and Becker, 2002). The comparison of miRNA expression between 5-aza-dC- and/or TSA-treated and parental cell lines is a feasible approach for the identification of differentially expressed cancer-related miRNAs. However, the involvement of the epigenetic regulation for miRNA expression is unlikely in at least the two NSCLC cell lines we studied. Recently, it was shown that the expression of miRNAs may be transcriptionally linked to the expression of other genes, coding for both proteins and noncoding RNAs (Rodriguez et al., 2004; Baskerville and Bartel, 2005). Indeed, approximately 30% of the 43

Table 3. Postoperative survival of patients with lung adenocarcinoma in relation to clinicopathological characteristics and miRNA expression analyzed by microarray analysis

Variable	Subset	Hazard ratio (95% confidence interval)	p
Univariate analysis (n = 65)			
Age	age ≥ 67/age < 67	1.41 (0.67–3.06)	0.348
Sex	male/female	1.36 (0.64–2.93)	0.413
Stage	II–IV/I	2.51 (1.29–6.82)	0.010
Smoking history	current/former	1.32 (0.63–2.79)	0.456
hsa-mir-155 (n = 55)	high/low	3.42 (1.42–8.19)	0.006
hsa-let-7a-2 (n = 52)	low/high	2.35 (1.08–6.86)	0.033
Multivariate analysis (n = 55)[a,b]			
Age	age ≥ 67/age < 67	1.92 (0.71–5.17)	0.195
Sex	male/female	1.23 (0.47–3.22)	0.669
Stage	II–IV/I	3.27 (1.31–8.37)	0.013
Smoking history	current/former	1.49 (0.51–4.34)	0.457
hsa-mir-155	high/low	3.03 (1.13–8.14)	0.027

[a]Multivariate analysis, Cox proportional hazard regression model.
[b]*hsa-let-7a-2* low/high was not statistically significant (p = 0.129).

miRNAs that showed different expression in lung cancer tissue versus noncancerous lung tissue are located within exons or introns of known protein-coding genes, such as *TMEM49* (transmembrane protein 49) for *hsa-mir-21*, *EGFL7* (EGF-like-domain, multiple 7) for *hsa-mir-126** and *hsa-mir-126*, *GABRE* (γ-aminobutyric acid [GABA] A receptor, epsilon) for *hsa-mir-224* and *SLIT3* (slit homolog 3) for *hsa-mir-218-2*. *TMEM49* is in a 17q23 amplicon that also contains the *PPM1D* (protein phosphatase 1D magnesium-dependent, delta isoform; *Wip1*) (Bulavin et al., 2002), which encodes Ser/Thr protein phosphatase, inactivates p53 tumor suppressor activity, and facilitates *RAS*- and *ERBB2*-induced murine mammary tumors (Bulavin et al., 2004). Increased *PPM1D* expression has not been previously detected in human lung cancer. *EGFL7* is expressed at high levels in the vasculature of proliferative tissue and is downregulated in mature vessels in the normal adult tissue (Parker et al., 2004). Future studies will determine the correlation between the expression of these exonic or intronic miRNAs and their host genes in lung cancers.

The global expression profile of miRNAs with Cox proportional hazard regression analysis could identify miRNAs that were associated with adenocarcinoma patient survival. The finding that expression of the five miRNAs (*hsa-mir-155*, *hsa-mir-17-3p*, *hsa-let-7a-2*, *hsa-mir-145*, and *hsa-mir-21*) is statistically altered in lung cancers and also has a prognostic impact on the survival warrants additional studies to investigate how altered miRNA expression would manifest the biological consequences in the development and/or progression of human cancers. It was recently reported that *hsa-mir-21* can function as an antiapoptotic factor in cultured glioblastoma cells (Chan et al., 2005). Because *hsa-mir-21* expression was upregulated significantly in lung cancer tissues, it was speculated that aberrant increased expression of the miRNA might block the expression of gene products that induce apoptosis and might be related to lung carcinogenesis. Interestingly, high *hsa-mir-155* expression had a significantly worse prognostic impact on adenocarcinoma patients as an independent risk factor and therefore could serve as a marker for survival. A unique 13 miRNA expression signature including *hsa-mir-155* was also a prognostic factor of chronic lymphocytic leukemia (Calin et al., 2005). Although *mir-155* is overexpressed in several types of human cancer, its biological function remains still uncertain. However, a previous study has shown that *BIC* (host gene of *hsa-mir-155*) is implicated as a collaborator with *c-myc* in an avian lymphoma model system (Tam et al., 2002). We were able to crossvalidate the clinical importance of outcome predictive miRNAs (*hsa-mir-155* and *hsa-let-7a-2*) using an independent additional case by real-time RT-PCR analysis. Again, a multivariate analysis revealed that high precursor *hsa-mir-155* expression independently contributed to patient outcome. In our study, only *hsa-let-7a-2* of the *let-7* family marginally correlated with prognosis in the original set of adenocarcinomas by miRNA microarray analysis. The *hsa-let-7a-2* expression data remained statistically sufficient in the original set of adenocarcinomas and was not statistically significant in the second independent set by real-time RT-PCR analysis. However, reduced *hsa-let-7a-2* expression correlated with poor survival by univariate analysis as well

Table 4. Postoperative survival of patients with lung adenocarcinoma in relation to clinicopathological characteristics and precursor miRNA expression analyzed by real-time RT-PCR analysis

		Original cohort (n = 32)		Additional cohort (n = 32)		All cases (n = 64)	
Variable	Subset	Hazard ratio (95% CI)	p	Hazard ratio (95% CI)	p	Hazard ratio (95% CI)	p
Univariate analysis							
Age	age ≥ 67/age < 67	1.89 (0.62–5.34)	0.274	1.21 (0.46–3.21)	0.679	1.28 (0.64–2.58)	0.482
Sex	male/female	0.53 (0.14–1.56)	0.232	1.37 (0.54–3.63)	0.479	0.99 (0.49–1.98)	0.975
Stage	II–IV/I	4.22* (1.91–23.6)	0.003	2.37* (1.01–7.83)	0.048	3.07* (1.82–8.84)	<0.001
Smoking history	current/former	0.92 (0.31–2.66)	0.921	1.22 (0.47–3.16)	0.674	1.12 (0.56–2.25)	0.757
precursor *hsa-mir-155*	high/low	2.75* (1.05–12.1)	0.047	2.52* (1.10–7.45)	0.033	2.74* (1.53–6.91)	0.002
precursor *hsa-let-7a-2*	low/high	3.01* (1.09–9.86)	0.037	2.22 (0.91–5.71)	0.084	2.73* (1.42–5.88)	0.003
Multivariate analysis							
Age	age ≥ 67/age < 67	0.91 (0.22–3.68)	0.899	0.93 (0.30–2.91)	0.914	1.22 (0.58–2.53)	0.593
Sex	male/female	0.35 (0.11–1.17)	0.089	0.92 (0.32–2.66)	0.885	0.85 (0.41–1.74)	0.659
Stage	II–IV/I	8.99* (1.95–41.2)	0.004	4.91* (1.51–15.9)	0.008	5.58* (2.42–12.8)	<0.001
Smoking history	current/former	1.01 (0.30–3.38)	0.980	2.27 (0.70–7.34)	0.170	1.89 (0.85–4.21)	0.117
precursor *hsa-mir-155*	high/low	13.3* (2.59–69.0)	0.002	3.77* (1.32–10.6)	0.013	4.98* (2.29–10.8)	<0.001
precursor *hsa-let-7a-2*	low/high	3.93* (1.06–14.5)	0.040	2.97* (1.07–8.23)	0.036	3.55* (1.64–7.69)	0.001

CI, confidence interval; multivariate analysis, Cox proportional hazard regression model. Asterisks indicate statistically significant values.

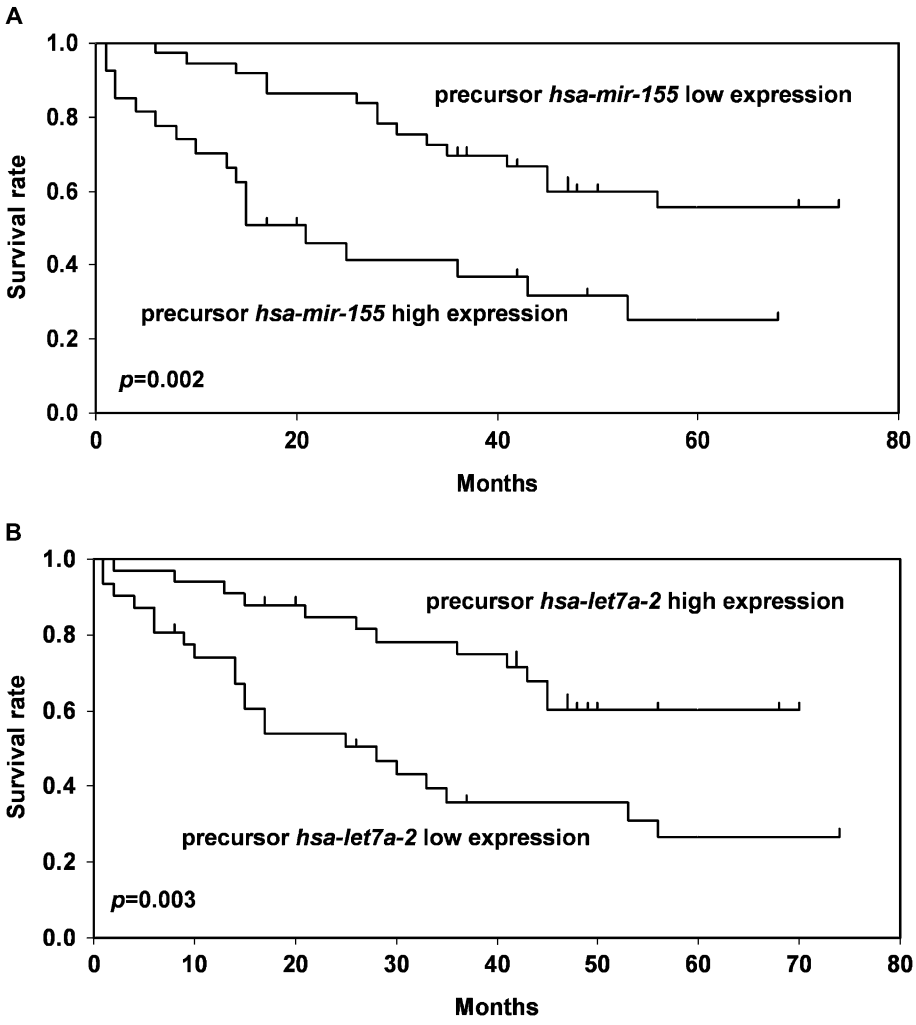

Figure 3. Kaplan-Meier survival curves for adenocarcinoma patients

Sixty-four adenocarcinoma cases from two combined independent cohorts were classified according to either precursor *hsa-mir-155* expression or precursor *hsa-let-7a-2* expression estimated by real-time RT-PCR analysis. The survival data were compared using the log-rank test.

A: Precursor *hsa-mir-155* high expression (n = 27), group with the expression ratio ≥ mean expression ratio (1.19). Precursor *hsa-mir-155* low expression (n = 37), group with the expression ratio < mean expression ratio.

B: Precursor *hsa-let-7a-2* high expression (n = 33), group with the expression ratio ≥ mean expression ratio (0.92). Precursor *hsa-let-7a-2* low expression (n = 31), group with the expression ratio < mean expression ratio.

as multivariate analysis in the combined set of two independent cohorts, suggesting that it is a prognostic factor in lung cancer, consistent with a previous report (Takamizawa et al., 2004; Johnson et al., 2005).

Several publications have presented algorithms with which to identify putative targets for miRNA (Lewis et al., 2003; John et al., 2004). However, the prediction and validation of target mRNAs by computerized means and experimental approaches is a still unresolved task. Recently, it was shown that the *let-7* family negatively regulates RAS in *C. elegans* as well as human cells, and the downregulation of *let-7* could result in the upregulation of RAS and induce oncogenesis in human lung cancer (Johnson et al., 2005). The GO analysis that we conducted for putative targets of *let-7a* was consistent with these findings, showing an association with target transcripts involved in the intracellular signaling. In addition, our GO analysis for *hsa-mir-155* suggests a role for this miRNA in regulating target transcripts associated with transcription. Besides this GO analysis, the web-based computational approaches to predict gene targets were performed for *hsa-mir-155* (miRBase Targets BETA Version 1.0, PicTar predictions, and TargetScan). Table S3 shows ten putative target genes that were commonly predicted by three different programs and indicates that the cancer-associated genes are potentially regulated by this miRNA. However, additional studies are needed to identify the targets of the miRNAs and to experimentally correlate them with lung carcinogenesis.

In conclusion, human lung cancer has extensive alterations of miRNA expression that may deregulate cancer-related genes. The miRNA molecular profiles of lung adenocarcinoma also correlate with patient survival.

Experimental procedures

Samples

One hundred and four pairs of primary lung cancers and corresponding noncancerous lung tissues were used in this study. An additional 32 cases, which could be followed up until 5 years, were used for an independent validation data set. These specimens were obtained from patients in the Baltimore metropolitan area from 1990 to 1999 with informed consent and IRB agreement. For the majority of samples, clinical and biological information was available. Total RNA were isolated by TRIzol (Invitrogen) according to the manufacturer's instructions.

Microarray analysis

Microarray analysis was performed as previously described (Liu et al., 2004). Briefly, 5 μg of total RNA was used for hybridization on miRNA microarray chips containing 352 probes in triplicate. The microarrays were hybridized in 6× SSPE (0.9 M NaCl/60 mM $NaH_2PO_4 \cdot 2H_2O$/8 mM EDTA [pH 7.4])/30% formamide at 25°C for 18 hr, washed in 0.75× TNT (Tris-HCl/NaCl/Tween 20) at 37°C for 40 min, and processed by using a method of direct

detection of the biotin-containing transcripts by streptavidin-Alexa 647 conjugate. Processed slides were scanned using a PerkinElmer ScanArray XL5K Scanner.

An average value of the three spot replicates of each miRNA was normalized and analyzed in BRB-ArrayTools version 3.2.3. After excluding negative values with hybridization intensity below background, normalization was performed by using the median normalization method and normalization to median array as reference. Finally, we selected 147 miRNAs with consistent log values present in more than 50% of the samples. We identified genes that were differently expressed among groups using Student's t tests or F tests, and genes were considered statistically significant if their p value was less than 0.001. We also performed a global test of whether the expression profiles differed between the groups by permuting the labels of which arrays corresponded to which groups. For each permutation, the p values were recomputed, and the number of genes significant at the 0.001 level was noted. The proportion of the permutations that gave at least as many significant genes as with the actual data was the significance level of the global test.

For phenotypical and histological comparisons, we performed the class prediction analysis based on the compound covariate predictor. We estimated the prediction error using leave-one-out crossvalidation. We also evaluated whether the crossvalidated error rate for a model was significantly less than one would expect from random prediction. The class labels were randomly permutated, and the entire leave-one-out crossvalidation process was repeated 2000 times. All data were submitted to the ArrayExpress database, and the accession number are E-TABE-22.

Solution hybridization detection analysis and real-time RT-PCR analysis

The expression levels of mature miRNAs were measured by solution hybridization detection method with *mir*Vana miRNA Detection Kit (Ambion Inc., TX). Briefly, total RNA (1 μg) was incubated with radiolabeled probes, which were prepared by 5′ end labeling by T4 Polynucleotide Kinase. Following digestion to remove the probe that was not bound by target miRNA, the radiolabeled products were fractionated by denaturing polyacrylamide gel electrophoresis.

Real-time RT-PCR analysis was performed as described (Schmittgen et al., 2004). Briefly, RNA was reverse transcribed to cDNA with gene-specific primers and Thermoscript, and the relative amount of each miRNA to tRNA for initiator methionine was described, using the equation 2^{-dCT}, where $dC_T = (C_{TmiRNA} - C_{TU6})$. The probe and primer sequences are available upon request.

5-aza-dC and/or TSA treatment

For the first 48 hr, A549 and NCI-H157 cells (American Tissue Culture Collection) were incubated with medium containing 1.0 μM 5-aza-dC (Sigma); the cells were then incubated for another 24 hr with the addition of 1.0 μM TSA (Sigma). Total RNA was isolated, and then microarray analysis was performed as described above. Each treatment was performed in triplicate.

Survival analysis

We identified genes whose expression was significantly related to survival of the patient. We computed a statistical significance level for each gene based on a univariate Cox proportional hazard regression model in BRB-ArrayTools version 3.2.3. These p values were then used in a multivariate permutation test in which the survival times and censoring indicators were randomly permuted among arrays. Genes were considered statistically significant if their p value was less than 0.05.

The survival curves were estimated by the Kaplan-Meier method, and the resulting curves were compared using the log-rank test. The joint effect of covariables was examined using the Cox proportional hazard regression model. Statistical analysis was performed using StatMate (ATMS Co., Ltd., Tokyo, Japan).

GO analysis

Predicted targets of *hsa-mir-155* and *hsa-let-7a* were determined by the methods of Lewis et al. (2005) and PicTar (Krek et al., 2005) and were analyzed with respect to the overrepresentation within different biological grouping categories including GO. Briefly, the predicted target gene lists were subjected to analysis using WholePathwayScope (WPS) (Yi et al., 2006). The level of overrepresentation is measured based on Fisher's exact test on a 2 × 2 contingency table for each GO term (whether a gene is in the given list or not versus whether this gene is associated with a GO term or not). Then, Fisher's exact test p values were computed for each term in each GO and ranked from smaller to higher values to estimate the statistical significance and priority for each term. Those terms with Fisher's exact test p values less than 0.005 were listed.

Supplemental data

The Supplemental Data include two supplemental figures and three supplemental tables and can be found with this article online at http://www.cancercell.org/cgi/content/full/9/3/189/DC1/.

Acknowledgments

We thank Dr. Xin Wang for helpful discussion, Dorothea Dudek and the NCI, CCR Fellows Editorial Board for editorial assistance, Drs. Krista Zanetti and Leah Mechanic for statistical analysis, and Judith A. Weish for help in figure printing. We also thank Drs. Raymond T. Jones, Andrew Borkowski, and Mark J. Krasna at the University of Maryland and Baltimore Veterans Administration for pathological diagnosis and sample collection as well as Audrey Salabes for interviewing. Analyses were performed using BRB-ArrayTools developed by Dr. Richard Simon and Amy Peng Lam. This research was supported by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. This work was supported in part by grant CA76259 to C.M.C. and by federal funds from the National Cancer Institute, National Institute of Health, under contract number NO1-CO-12400 to R.M.S. N.Y. is supported by the Uehara Memorial Foundation of Japan.

Received: June 10, 2005
Revised: October 28, 2005
Accepted: January 23, 2006
Published: March 13, 2006

References

Ambros, V. (2003). MicroRNA pathways in flies and worms: Growth, death, fat, stress, and timing. Cell *113*, 673–676.

Baskerville, S., and Bartel, D.P. (2005). Microarray profiling of microRNAs reveals frequent coexpression with neighboring miRNAs and host genes. RNA *11*, 241–247.

Bulavin, D.V., Demidov, O.N., Saito, S., Kauraniemi, P., Phillips, C., Amundson, S.A., Ambrosino, C., Sauter, G., Nebreda, A.R., Anderson, C.W., et al. (2002). Amplification of PPM1D in human tumors abrogates p53 tumor-suppressor activity. Nat. Genet. *31*, 210–215.

Bulavin, D.V., Phillips, C., Nannenga, B., Timofeev, O., Donehower, L.A., Anderson, C.W., Appella, E., and Fornace, A.J., Jr. (2004). Inactivation of the Wip1 phosphatase inhibits mammary tumorigenesis through p38 MAPK-mediated activation of the p16(Ink4a)-p19(Arf) pathway. Nat. Genet. *36*, 343–350.

Calin, G.A., Dumitru, C.D., Shimizu, M., Bichi, R., Zupo, S., Noch, E., Aldler, H., Rattan, S., Keating, M., Rai, K., et al. (2002). Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc. Natl. Acad. Sci. USA *99*, 15524–15529.

Calin, G.A., Liu, C.G., Sevignani, C., Ferracin, M., Felli, N., Dumitru, C.D., Shimizu, M., Cimmino, A., Zupo, S., Dono, M., et al. (2004a). MicroRNA profiling reveals distinct signatures in B cell chronic lymphocytic leukemias. Proc. Natl. Acad. Sci. USA *101*, 11755–11760.

Calin, G.A., Sevignani, C., Dumitru, C.D., Hyslop, T., Noch, E., Yendamuri, S., Shimizu, M., Rattan, S., Bullrich, F., Negrini, M., and Croce, C.M. (2004b). Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc. Natl. Acad. Sci. USA *101*, 2999–3004.

Calin, G.A., Ferracin, M., Cimmino, A., Di Leva, G., Shimizu, M., Wojcik, S.E., Iorio, M.V., Visone, R., Sever, N.I., Fabbri, M., et al. (2005). A microRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N. Engl. J. Med. *353*, 1793–1801.

Chan, J.A., Krichevsky, A.M., and Kosik, K.S. (2005). MicroRNA-21 is an antiapoptotic factor in human glioblastoma cells. Cancer Res. *65*, 6029–6033.

Eberharter, A., and Becker, P.B. (2002). Histone acetylation: a switch between repressive and permissive chromatin. Second in review series on chromatin dynamics. EMBO Rep. *3*, 224–229.

Eis, P.S., Tam, W., Sun, L., Chadburn, A., Li, Z., Gomez, M.F., Lund, E., and Dahlberg, J.E. (2005). Accumulation of miR-155 and BIC RNA in human B cell lymphomas. Proc. Natl. Acad. Sci. USA *102*, 3627–3632.

Granville, C.A., and Dennis, P.A. (2005). An overview of lung cancer genomics and proteomics. Am. J. Respir. Cell Mol. Biol. *32*, 169–176.

Gregory, R.I., and Shiekhattar, R. (2005). MicroRNA biogenesis and cancer. Cancer Res. *65*, 3509–3512.

Hayashita, Y., Osada, H., Tatematsu, Y., Yamada, H., Yanagisawa, K., Tomida, S., Yatabe, Y., Kawahara, K., Sekido, Y., and Takahashi, T. (2005). A polycistronic microRNA cluster, miR-17-92, is overexpressed in human lung cancers and enhances cell proliferation. Cancer Res. *65*, 9628–9632.

He, L., Thomson, J.M., Hemann, M.T., Hernando-Monge, E., Mu, D., Goodson, S., Powers, S., Cordon-Cardo, C., Lowe, S.W., Hannon, G.J., and Hammond, S.M. (2005). A microRNA polycistron as a potential human oncogene. Nature *435*, 828–833.

Hutvagner, G., and Zamore, P.D. (2002). A microRNA in a multiple-turnover RNAi enzyme complex. Science *297*, 2056–2060.

Iorio, M.V., Ferracin, M., Liu, C.G., Veronese, A., Spizzo, R., Sabbioni, S., Magri, E., Pedriali, M., Fabbri, M., Campiglio, M., et al. (2005). MicroRNA gene expression deregulation in human breast cancer. Cancer Res. *65*, 7065–7070.

John, B., Enright, A.J., Aravin, A., Tuschl, T., Sander, C., and Marks, D.S. (2004). Human microRNA targets. PLoS Biol. *2*, e363. 10.1371/journal.pbio.0020363.

Johnson, S.M., Grosshans, H., Shingara, J., Byrom, M., Jarvis, R., Cheng, A., Labourier, E., Reinert, K.L., Brown, D., and Slack, F.J. (2005). RAS is regulated by the let-7 microRNA family. Cell *120*, 635–647.

Jones, P.A., and Baylin, S.B. (2002). The fundamental role of epigenetic events in cancer. Nat. Rev. Genet. *3*, 415–428.

Karube, Y., Tanaka, H., Osada, H., Tomida, S., Tatematsu, Y., Yanagisawa, K., Yatabe, Y., Takamizawa, J., Miyoshi, S., Mitsudomi, T., and Takahashi, T. (2005). Reduced expression of Dicer associated with poor prognosis in lung cancer patients. Cancer Sci. *96*, 111–115.

Krek, A., Grun, D., Poy, M.N., Wolf, R., Rosenberg, L., Epstein, E.J., MacMenamin, P., da Piedade, I., Gunsalus, K.C., Stoffel, M., and Rajewsky, N. (2005). Combinatorial microRNA target predictions. Nat. Genet. *37*, 495–500.

Lagos-Quintana, M., Rauhut, R., Lendeckel, W., and Tuschl, T. (2001). Identification of novel genes coding for small expressed RNAs. Science *294*, 853–858.

Lau, N.C., Lim, L.P., Weinstein, E.G., and Bartel, D.P. (2001). An abundant class of tiny RNAs with probable regulatory roles in Caenorhabditis elegans. Science *294*, 858–862.

Lee, R.C., and Ambros, V. (2001). An extensive class of small RNAs in Caenorhabditis elegans. Science *294*, 862–864.

Lee, Y., Jeon, K., Lee, J.T., Kim, S., and Kim, V.N. (2002). MicroRNA maturation: stepwise processing and subcellular localization. EMBO J. *21*, 4663–4670.

Lee, Y., Ahn, C., Han, J., Choi, H., Kim, J., Yim, J., Lee, J., Provost, P., Radmark, O., Kim, S., and Kim, V.N. (2003). The nuclear RNase III Drosha initiates microRNA processing. Nature *425*, 415–419.

Lewis, B.P., Shih, I.H., Jones-Rhoades, M.W., Bartel, D.P., and Burge, C.B. (2003). Prediction of mammalian microRNA targets. Cell *115*, 787–798.

Lewis, B.P., Burge, C.B., and Bartel, D.P. (2005). Conserved seed pairing, often flanked by adenosines, indicates that thousands of human genes are microRNA targets. Cell *120*, 15–20.

Liu, C.G., Calin, G.A., Meloon, B., Gamliel, N., Sevignani, C., Ferracin, M., Dumitru, C.D., Shimizu, M., Zupo, S., Dono, M., et al. (2004). An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc. Natl. Acad. Sci. USA *101*, 9740–9744.

Lu, J., Getz, G., Miska, E.A., Alvarez-Saavedra, E., Lamb, J., Peck, D., Sweet-Cordero, A., Ebert, B.L., Mak, R.H., Ferrando, A.A., et al. (2005). MicroRNA expression profiles classify human cancers. Nature *435*, 834–838.

McManus, M.T. (2003). MicroRNAs and cancer. Semin. Cancer Biol. *13*, 253–258.

Metzler, M., Wilda, M., Busch, K., Viehmann, S., and Borkhardt, A. (2004). High expression of precursor microRNA-155/BIC RNA in children with Burkitt lymphoma. Genes Chromosomes Cancer *39*, 167–169.

Meyerson, M., and Carbone, D. (2005). Genomic and proteomic profiling of lung cancers: lung cancer classification in the age of targeted therapy. J. Clin. Oncol. *23*, 3219–3226.

Michael, M.Z., O'Connor, S.M., van Holst Pellekaan, N.G., Young, G.P., and James, R.J. (2003). Reduced accumulation of specific microRNAs in colorectal neoplasia. Mol. Cancer Res. *1*, 882–891.

Minna, J.D., Roth, J.A., and Gazdar, A.F. (2002). Focus on lung cancer. Cancer Cell *1*, 49–52.

Parker, L.H., Schmidt, M., Jin, S.W., Gray, A.M., Beis, D., Pham, T., Frantz, G., Palmieri, S., Hillan, K., Stainier, D.Y., et al. (2004). The endothelial-cell-derived secreted factor Egfl7 regulates vascular tube formation. Nature *428*, 754–758.

Rodriguez, A., Griffiths-Jones, S., Ashurst, J.L., and Bradley, A. (2004). Identification of mammalian microRNA host genes and transcription units. Genome Res. *14*, 1902–1910.

Schmittgen, T.D., Jiang, J., Liu, Q., and Yang, L. (2004). A high-throughput method to monitor the expression of microRNA precursors. Nucleic Acids Res. *32*, e43.

Sekido, Y., Fong, K.M., and Minna, J.D. (2003). Molecular genetics of lung cancer. Annu. Rev. Med. *54*, 73–87.

Storey, J.D., and Tibshirani, R. (2003). Statistical significance for genome-wide studies. Proc. Natl. Acad. Sci. USA *100*, 9440–9445.

Takamizawa, J., Konishi, H., Yanagisawa, K., Tomida, S., Osada, H., Endoh, H., Harano, T., Yatabe, Y., Nagino, M., Nimura, Y., et al. (2004). Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res. *64*, 3753–3756.

Tam, W., Hughes, S.H., Hayward, W.S., and Besmer, P. (2002). Avian bic, a gene isolated from a common retroviral site in avian leukosis virus-induced lymphomas that encodes a noncoding RNA, cooperates with c-myc in lymphomagenesis and erythroleukemogenesis. J. Virol. *76*, 4275–4286.

W.D. Travis, E. Brambilla, C.C. Harris, and H.K. Muller-Hermelink, eds. (2004). World Health Organization Classification of Tumours, Pathology and Genetics: Tumours of the Lung, Pleura, Thymus and Heart (Lyon: IARC Press), pp. 12–15.

Yi, R., Qin, Y., Macara, I.G., and Cullen, B.R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. *17*, 3011–3016.

Yi, M., Horton, J.D., Cohen, J.C., Hobbs, H.H., and Stephens, R.M. (2006). WholePathwayScope: a comprehensive pathway-based analysis tool for high-throughput data. BMC Bioinformatics *7*, 30.

Yokota, J., and Kohno, T. (2004). Molecular footprints of human lung cancer progression. Cancer Sci. *95*, 197–204.

Volinia, S., Calin, G.A., Liu, C.G., Ambs, S., Cimmino, A., Petrocca, F., Visone, R., Iorio, M., Roldo, C., Ferracin, M., et al. (2006). A microRNA expression signature of human solid tumors defines cancer gene targets. Proc. Natl. Acad. Sci. USA *103*, 2257–2261.

Accession numbers

All microarray data were submitted to the ArrayExpress database, and the accession number is E-TABE-22.