Distinctive MicroRNA Profiles Relating to Patient Survival in Esophageal Squamous Cell Carcinoma

Yong Guo,[1,5,6] Zhaoli Chen,[3] Liang Zhang,[5,6] Fang Zhou,[3] Susheng Shi,[6] Xiaoli Feng,[4] Baozhong Li,[3] Xin Meng,[5,6] Xi Ma,[5,6] Mingyong Luo,[5] Kang Shao,[3] Ning Li,[3] Bin Qiu,[3] Keith Mitchelson,[1,5,6] Jing Cheng,[1,2,5,6] and Jie He[3]

[1]Medical Systems Biology Research Center and [2]The State Key Laboratory of Biomembrane and Membrane Biotechnology, Tsinghua University; Departments of [3]Thoracic Surgery and [4]Pathology, Cancer Hospital and Institute, Peking Union Medical College and Chinese Academy of Medical Sciences, Beijing, China; and [5]National Engineering Research Center for Beijing Biochip Technology and [6]CapitalBio Corporation, Changping District, Beijing, China

Abstract

Esophageal cancer is the sixth leading cause of death from cancer and one of the least studied cancers worldwide. The global microRNA expression profile of esophageal cancer has not been reported previously. Here, for the first time, we have investigated expressed microRNAs in cryopreserved esophageal cancer tissues using advanced microRNA microarray techniques. Our microarray analyses identified seven microRNAs that could distinguish malignant esophageal cancer lesions from adjacent normal tissues. Some microRNAs could be correlated with the different clinicopathologic classifications. High expression of hsa-miR-103/107 correlated with poor survival by univariate analysis as well as by multivariate analysis. These results indicate that microRNA expression profiles are important diagnostic and prognostic markers of esophageal cancer, which might be analyzed simply using economical approaches such as reverse transcription-PCR. [Cancer Res 2008;68(1):26–34]

Introduction

Human esophageal cancer occurs worldwide with a variable geographic distribution and ranks eighth in order of occurrence and sixth as the leading cause of cancer mortality, affecting men more than women (1). A 20-fold variation is observed in its incidence between low-risk western Africa and high-risk northern China where it exceeds 100 in 100,000 individuals. It has two main forms, each with distinct etiologic and pathologic characteristics, esophageal squamous cell carcinoma (ESCC) and adenocarcinoma. ESCC is the most frequent subtype of esophageal cancer, although the incidence of adenocarcinoma in the western world is increasing faster than other esophageal malignancies. At diagnosis, nearly 50% of patients have cancer that extends beyond the primary locoregional confines, and ∼75% of patients requiring surgery have proximal lymph node metastases. Although tumor-node-metastasis (TNM) classification allows diagnosis of the tumor, it provides little therapeutic biological information, such as the metastatic potential or the sensitivity or resistance of the tumor to radiotherapy and chemotherapy. There is an urgent need for accurate prognostic indicators to distinguish high-risk patients from other patients, so that optimal treatments can be designed. Research over the last 20 years has identified a number of oncogenic and tumor-suppressor proteins that are associated with induction of ESCC (2–8), yet molecular indicators of the origin of cellular deregulation in ESCC have not been identified.

MicroRNAs (miRNAs) are a species of small noncoding single-stranded RNA of about 21 to 23 nucleotides that through partial sequence homology may interact with the 3′-untranslated region of target mRNA molecules (9). Growing evidence has indicated important roles for different miRNA species in the development of different cancers (10–17). Recently, genome-wide expression of miRNAs has been able to be examined by microarray (18–23) and on a more limited miRNA set by microbead hybridization (24) or reverse transcription-PCR (RT-PCR; ref 25). These miRNA expression studies have confirmed the initially observed deregulation of individual miRNAs and have identified changes in the pattern of expression of a large number of miRNAs in individual cancers. Notwithstanding the keen interest in miRNA expression in cancers, the global expression of miRNA in human esophageal cancer has not been determined previously.

Extensive collections of archived tissue biopsies and paired samples exist in hospital laboratories and biobanks, frequently with extensive clinicopathologic information and disease outcomes. However, archived samples may have a variable quality of preserved biomolecules, often displaying poor RNA preservation (26). Despite these constraints, the development of methods to investigate stored samples has proved essential as molecular analysis can provide initial indications of altered biochemical pathways and provide biomarkers that correlate with patient survival and other clinical variables and, thus, aid the diagnosis of cancer and the prognosis for management of the disease (27, 28).

Here, we present the results of a genome-wide miRNA expression profiling in paired sets of frozen archival tissues from the esophageal tumors and corresponding adjacent normal tissues. We observed unique miRNA expression signatures that could distinguish malignant from adjacent normal esophageal tissues. We found that the expression of particular miRNAs was altered in human esophageal cancer and determined that particular miRNA profiles correlate closely with patient survival and with several clinicopathologic indicators. We have also compared the miRNA profiles of the stored tissues to a set of fresh tissues and could show a similarity in the miRNA profiles of archival and fresh materials.

Materials and Methods

Patients and samples. Thirty-one pairs of primary esophageal squamous cell cancer tissues and corresponding adjacent normal esophageal tissues were used as a training group. These specimens were obtained from patients in the Cancer Institute and Hospital, Chinese Academy of Medical Sciences (CAMS) from 1999 to 2001 with informed consent and agreement. All tissue samples were from untreated patients undergoing surgery and were snap frozen in liquid nitrogen and stored (minimum of 5 years) at −80°C until the extraction of RNA. A second independent set of fresh tissues from 24 paired samples and 1 unpaired cancer sample was used as an independent validation data set, collected from the same hospital in 2006, and stored for <6 months. Samples from a further 22 cases of ESCC with follow-up information (minimum of 5 years) were used for an independent validation of survival analysis. Peripheral portions of the resected esophageal samples were paraffin embedded, sectioned, and H&E stained using routine methods. The tumor cell concentrations were evaluated, and the tumor histology was independently confirmed by two pathologists. Follow-up information was extracted from the follow-up registry of the Cancer Institute and Hospital, CAMS. For all the samples, clinicopathologic information (age, gender, pathology, differentiation, TNM classification, tumor stage, and survival time after surgery) was available. The study was approved by the medical ethics committee of Cancer Institute and Hospital, CAMS.

Fabrication of the miRNA microarray. Altogether, 509 mature miRNA sequences were assembled and integrated into our miRNA microarray design. These comprised 435 human miRNAs including a further 122 predicted miRNA sequences from published references (29) and some 196 rat and 261 mouse mature miRNAs from the miRNA Registry.[7] In addition, we designed eight short oligonucleotides that possessed no homology to any known RNA sequence and generated their corresponding synthetic miRNAs by *in vitro* transcription using the Ambion miRNA Probe Construction kit (Cat. No.1550). Various amounts of these synthetic miRNAs were added into the human miRNA samples before analysis to act as external controls.

All of the miRNA probe sequences was designed to be fully complementary to their cognate mature miRNA. To facilitate probe immobilization onto the aldehyde-modified surface of the glass slides (CapitalBio), the probe sequences were concatenated up to a length of 40 nt (3′-end miRNA probe plus 5′-end polyT) and attached to the activated slide surface via a C6 5′-amino-modifier. Oligonucleotide probes were synthesized at MWG Biotech and dissolved in EasyArray spotting solution (CapitalBio) at a concentration of 40 μmol/L. Each probe was printed in triplicate using a SmartArray-136 microarrayer (CapitalBio).

Labeling of target RNAs. Total RNA was extracted with TRIZOL reagent (Invitrogen), and the low-molecular-weight RNA was isolated using a PEG solution precipitation method (30). The low-molecular-weight RNA was labeled using the T4 RNA ligase labeling method described by Thomson et al. (19). In brief, 4 μg of low-molecular-weight RNA were labeled with 500 ng of 5′-phosphate-cytidyl-uridyl-cy3-3′ (Dharmacon) with 2 units of T4 RNA ligase (New England Biolabs). The labeling reaction was performed at 4°C for 2 h. Labeled RNA was precipitated with 0.3 mol/L sodium acetate and 2.5 volumes ethanol, and after washing with ethanol and drying, it was resuspended in 15 μL of hybridization buffer containing 3 × SSC, 0.2% SDS, and 15% formamide.

Slide hybridization. Hybridization was performed at 42°C under LifterSlip (Erie) in a hybridization cassette which was placed in a three-dimensional–tilting agitator BioMixer II (CapitalBio) to provide continuous mixing of the hybridization buffer and more uniform hybridization across the entire slide surface, preventing edge effects and giving improved signal intensity. The efficiency of these measures has been shown previously in genome-wide mRNA expression profiling studies (31). The array was then washed with two consecutive washing solutions of 0.2% SDS, 2 × SSC at 42°C for 5 min, and 0.2% SSC for 5 min at room temperature. Arrays were

scanned with a LuxScan 10K-A laser confocal scanner, and the images obtained were then analyzed using LuxScan 3.0 software (both from CapitalBio).

Computational analysis. For all samples, after average values of the replicate spots of each miRNA were background subtracted, faint spots were filtered out if the expression signal was <800. Signals were normalized using the median center tool for genes in the Cluster 3.0 software before performing the unsupervised clustering with complete linkage and uncentered Pearson correlation to reveal the underlying structure of the miRNA expression (32). Differentially expressed miRNAs were identified by significance analysis of microarrays (SAM; ref. 33).[8] We progressively eliminated the miRNA with the lowest score(d) deduced by SAM (in the previous cycle) and used the remaining miRNAs to build a model and estimate its accuracy, until all miRNAs were eliminated. The miRNA set with highest accuracy was considered as the sufficient minimum marker set. To establish a classifier, we used seven strategies for dimensionality reduction [no extraction; first and second; first and third; first, second, and third components by principal component analysis (PCA); first and second; first and third; first, second, and third components by partial least squares method analyses; ref. 34] and six strategies for model building [linear support vector machine (SVM), one neighbor k-nearest neighbors (KNN), three neighbors KNN, five neighbors KNN, linear discriminate analysis, and quadratic discriminate analysis; ref. 35]. The 632 Bootstrap method was used to estimate the accuracy of each predicted model for the original training set (31 paired samples) by using random resampling with replacement >1,000 independent analyses (36). The accuracy was calculated using the formula, $acc_{boot} = \frac{1}{n} \sum_{i=1}^{n} (0.368 \times acc_{train}^{i} + 0.632 \times acc_{test}^{i})$, where n is the number of repeats, acc_{train} is the i-th experiment train accuracy and acc_{train}^{i} is the i-th experiment test accuracy.

The most significant predicted miRNA targets were analyzed by using four publicly available algorithms, i.e., miRBase,[9] MIRANDA,[10] TARGETSCAN,[11] and PICTAR.[12] To reduce the number of false positives, only putative target genes predicted by at least three of the programs were accepted. Patient survival curves were estimated by the Kaplan-Meier method. The joint effect of covariables was examined by using the Cox proportional hazard regression model. The gene functions were annotated by using the Gene Ontology, Biocarta, KEGG, and GenMAPP databases. All miRNA expression data have been submitted to the Gene Expression Omnibus[13] with the series accession number GSE 6188.

Quantitative RT-PCR analysis. For verification of miRNA expression profiles, total cellular RNAs were subjected to quantitative RT-PCR (qRT-PCR) with microRNA specific primers. Reverse transcriptase reactions contained 2.5 ng/μL total RNA, 25 nmol/L stem-loop reverse transcription (RT) primer, 1× RT buffer, 0.25 mmol/L each of deoxynucleotide triphosphates, 200 units M-MLV reverse transcriptase, and 0.25 units/mL RNase inhibitor (Invitrogen). The 7.5-μL reactions were incubated in an MJ Research PTC-225 Thermocycler for 30 min at 16°C, 30 min at 42°C, 5 min at 85°C, and then held at 4°C. All reverse transcriptase reactions, including no-template controls, were run in duplicate. qRT-PCRs were performed as previously described (25) with the following modifications. A FastStart DNA Master SYBR green I kit and a LightCycler (both from Roche Diagnostics) were used, following the manufacturer's protocols. The 10-μL PCR reaction contained 1 μL RT product, 1× PCR Master Mix, 15 nmol/L forward primer, and 15 nmol/L reverse primer. The reactions were incubated at 95°C for 10 min, followed by 40 cycles of 95°C for 15 s, 60°C for 35 s, and 72°C for 3 s. All quantitative PCR reactions, including no-template controls, were performed in triplicate. The relative expression ratios of miRNAs were determined with the crossing point as the cycle number. The highly conserved and universally expressed small nRNA U6 was used as an

---

[7] http://microrna.sanger.ac.uk, miRBase release 7.0, accessed Sept. 2005.

[8] http://www-stat.stanford.edu/~tibs/SAM/index.html
[9] http://microrna.sanger.ac.uk/sequences/
[10] http://www.microrna.org/
[11] http://www.targetscan.org/
[12] http://pictar.bio.nyu.edu/
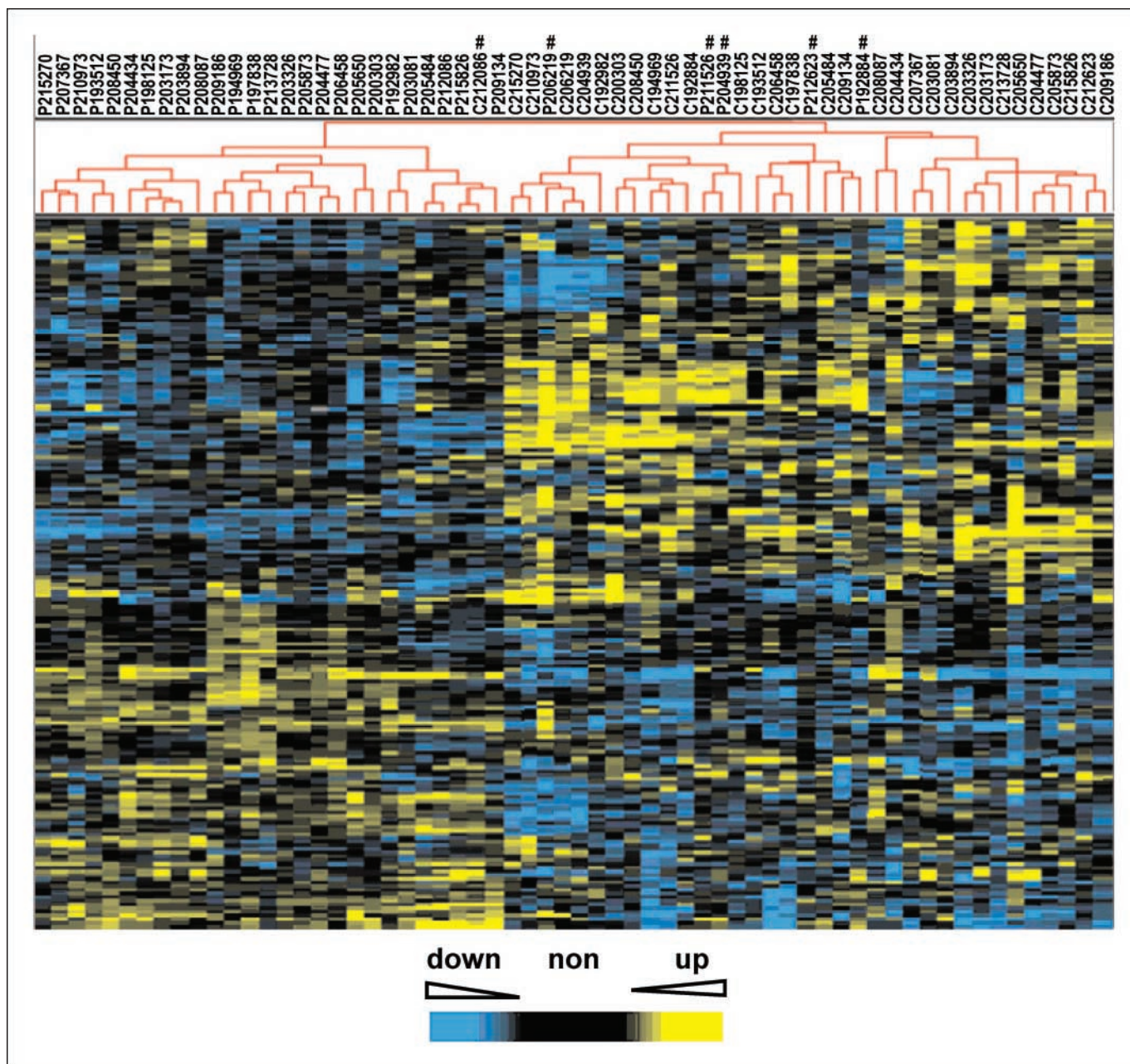[13] http://www.ncbi.nlm.nih.gov/geo

Figure 1. Unsupervised hierarchical clustering of miRNA expression. miRNA profiles of 31 paired samples from esophagus tissues were clustered using Cluster 3.0. The 31 paired samples are in columns and the 191 miRNAs are in rows. *C,* cancerous tissue; *P,* adjacent normal tissue. Nearly all cancerous and adjacent normal tissues were separated into two groups, except for C212086, P206219, P211526, P204939, P212623, and P192884. The samples labeled with # were wrongly classified pathologically.

endogenous control in the qRT-PCR. The results were analyzed using LightCycler software version 3.5 (Roche Diagnostics). The qRT-PCR amplification products were analyzed by melting curve analysis and confirmed by agarose gel electrophoresis.

Results

Altered miRNA expression in esophageal cancers and the identification of miRNAs associated with clinical features and disease progression. We analyzed the miRNA expression in 31 pairs of esophageal cancers and their corresponding adjacent normal tissues collected a minimum of 5 cm from the tumor. These tissues were initially snap frozen in liquid nitrogen and then stored

frozen at $-80°C$ for a minimum of 5 years until analysis. It was observed by formaldehyde gel electrophoresis analysis that the total RNA extracted from the cryoarchive-preserved tissues was extensively degraded, whereas the total RNA extracted from fresh tissues showed no such degradation. The expression signals of miRNAs from the fresh and archival tissues were compared and found to have nearly identical signal profiles (data not shown). We initially compared the expression intensity of the 191 miRNAs with signals detected above our defined signal threshold from all frozen samples. Comparison between all individual cancers and adjacent normal samples by complete linkage and uncentered Pearson correlation generated a hierarchical clustering of the samples on

the basis of similarity in the expression of any pairs of sample (Fig. 1). This initial unsupervised clustering successfully separated the 62 samples of cancerous tissues and adjacent normal tissues into 2 discrete groups, with the exception of 1 cancer sample and 5 adjacent normal samples.

Next, we asked whether the microarray data revealed specific molecular signatures for subsets of ESCC that differ in clinicopathologic classifications. We compared the miRNA expression of seven group pairs, including age, gender, gross pathologic classification, differentiation classification, different tumor stage classifications, and the entire cohort of pairs, as listed in Table 1. We used methods based on SAM tools for the two analyses, the first based directly on the miRNA signal strength in cancer tissues and the second on the ratio of the miRNA signals of cancer tissues versus paired adjacent normal tissues. Typically, many more miRNA genes were identified by direct signal strength than by ratio, yet several miRNAs were identified by both methods, with five miRNAs (hsa-miR-335, hsa-miR-181d, hsa-miR-25, hsa-miR-7, and hsa-miR-495) correlating with gross pathologic classification (fungating versus medullary) and two miRNAs (hsa-miR-25 and hsa-miR-130b) correlating with differentiation classification (high versus middle versus low). Tobacco and alcohol consumption are very strong risk factors for ESCC (37), yet when we compared the miRNA expression of pair groups, we found no miRNA related with either of these risk factors. Similarly, no miRNA expression correlated with age classification in our data set.

Establishment of a classifier to distinguish malignant esophageal tissues from normal tissues. First, 46 miRNAs were chosen from the training data (31 paired samples) by SAM with false discovery rate (FDR) is equals to 0. Subsequently, two steps were used to create a model: feature extraction and model building. As described in Materials and Methods, seven component-extraction strategies and six model-building methods were used. In total, 42 strategies were used to analyze the 46 miRNAs, and the best result of 96.97% was observed when the $PCA_{1,3}$-SVM strategy was used with a set of 7 miRNAs with the highest scored values (Fig. 2A), providing the definition of a classifier. Among the seven miRNAs, three miRNAs (hsa-miR-25, hsa-miR-424, and hsa-miR-151) showed up-regulation and four miRNAs (hsa-miR-100, hsa-miR-99a, hsa-miR-29c, and mmu-miR-140*) showed reduction in cancer versus normal tissue. After establishing the classifier, we used this model to assess the original training set (31 paired samples) along with an independent validation cohort (24 paired fresh samples and 1 unpaired fresh cancer sample) to confirm this modeling strategy. These analyses resulted in an accuracy of 98.38% with one misclassified sample in the original training set and an accuracy of 93.89% with a misclassification of three samples in the validation cohort (Fig. 2B). Overall, the results indicate that our strategy of model building was efficient and could readily distinguish malignant from normal esophageal tissues with as few as seven markers.

Correlation between miRNA expression profiles and prognosis of esophageal cancer patients. The median miRNA intensity value of the initial set of 31 patient samples (training cohort) was used as the cut-point in Kaplan-Meier survival analysis. The two mature forms of hsa-miR-103 and hsa-miR-107 are nearly

Table 1. Comparative analysis of clinicopathologic classifications

Classification (number)	miRNA selection based on signal strength*	miRNA selection based on signal ratio of cancer/normal[†]	FDR
Age classification			
Age < 49 (4) versus 49 ≤ Age ≥ 67 (20) versus Age > 67 (7)	0	0	0
Gender classification			
Male (23) versus female (8)	4 (hsa-miR-20a, -106a, 17.5p, 19b)	0	0
Gross pathologic classification			
Fungating (6) versus medullary (22)	28 (hsa-miR-375, 188, 95, 451, *7*, 200a, 429, 141, 19b, 19a, *495*, 10b, 148a, 15a, *335*, 130b, *25*, 186, 181c, *181d*, 491, and 146a; hsa-let-7i, predicted-mir88, 112, and 206; rno-miR-7 and mmu-miR-291.5p)	5 (hsa-miR-*335, 181d, 25, 7,* and *495*)	0
Differentiation classification			
High (10) versus middle (15) versus low (5)	5 (hsa-miR-335, *25, 130b*, 130a, 10a)	3 (hsa-miR-*25, 130b,* 181d,)	0
Tumor stage (T) classification			
T_1 (2) versus T_2 (3) versus T_3 (26)	0	1 (predicted miR-144)	0
Tumor stage (N) classification			
N_0 (16) versus N_1 (15)	13 (hsa-miR-199a*, 199a, 29b, 365, 146b, 155, 193a, 126, 23a, 195, 29a, 1, and 125a)	1 (has-miR-92)	0
Tumor stage (TNM) classification			
I (2) versus II (18) versus III (11)	23 (hsa-miR-146b, 193a, 130a, 106b, 142.5p, 19b, 30d, 30a.5p, 320, 195, 28, 125, 1, 26a, 126, 30c, 30b, 26b, and 23b; hsa-let-7a and 7d)	2 (predicted-mir144; rno-miR-347)	0.16/0

NOTE: The miRNAs selected by both methods are italicized.

*Each miRNA was selected based on miRNA signal strength in cancer tissues.

[†]miRNA selection based on signal ratio of cancer/normal: Each miRNA was selected based on the ratio of the miRNA signal in cancer tissues versus the signal in paired adjacent normal tissues.
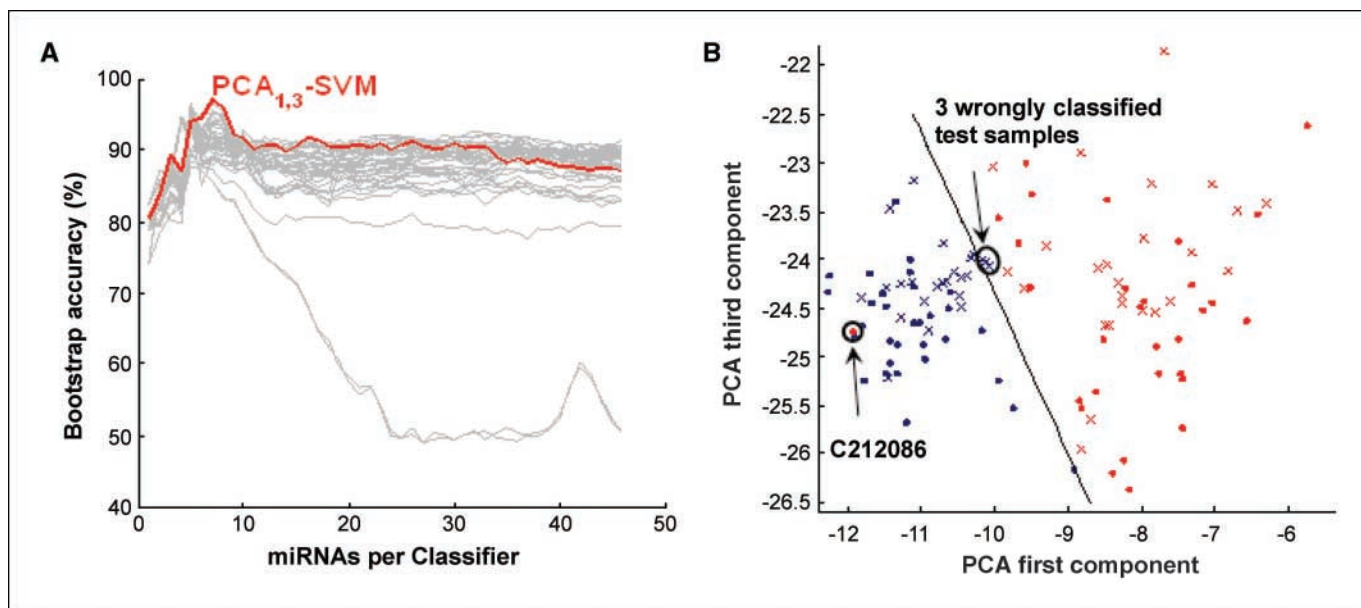
Figure 2. Establishment of a classifier to distinguish malignant esophageal tissues from normal tissues. *A,* estimated classifier accuracy versus number of miRNAs for different modeling strategies. A total of 42 combined strategies were performed, including some 7 strategies for feature extraction and 6 strategies for model building. The best result was gained by PCA$_{1,3}$-SVM (dimensionality reduction using first and third component by PCA, model building using SVM) and is indicated in red. *B,* PCA$_{1,3}$-SVM analysis. *Black line,* hyper plane built by SVM. *Red symbols,* cancer samples; *blue symbols,* adjacent normal samples. *Dot points,* training data set (31 paired long-preserved samples); *crosses,* represent the independent test data (24 paired fresh samples and 1 unpaired fresh cancer sample). One cancer sample was misclassified in the training data and three adjacent normal samples were misclassified in the test data. Training sample C212086 displayed significant error (*circled dot*) in the training data.

identical. The high similarity of the signals detected for both miRNAs is strongly suggestive of cross detection. For analysis purposes, they are treated here as a composite of both miRNAs. Here, hsa-miR-103/107 showed a strong correlation between low expression and high overall survival period (Fig. 3*A*). An additional 22 cases (test cohort) were analyzed for independent validation (Fig. 3*B*). The difference in the overall survival was statistically significant for hsa-miR-103/107 ($P = 0.013$, for training cohort;

$P = 0.041$, for test cohort; log-rank test). Kaplan-Meier survival analysis of the disease-free survival of patients in the training cohort gave a similar result (Supplementary Fig. S1). The disease-free survival analysis in the test cohort was not performed, because the disease-free information for most of these new cases was not available. Univariate Cox analysis in training and test cohorts (Supplementary Table S1) and in all 52 investigated patients (training set plus test set) with hsa-miR-103/107 and
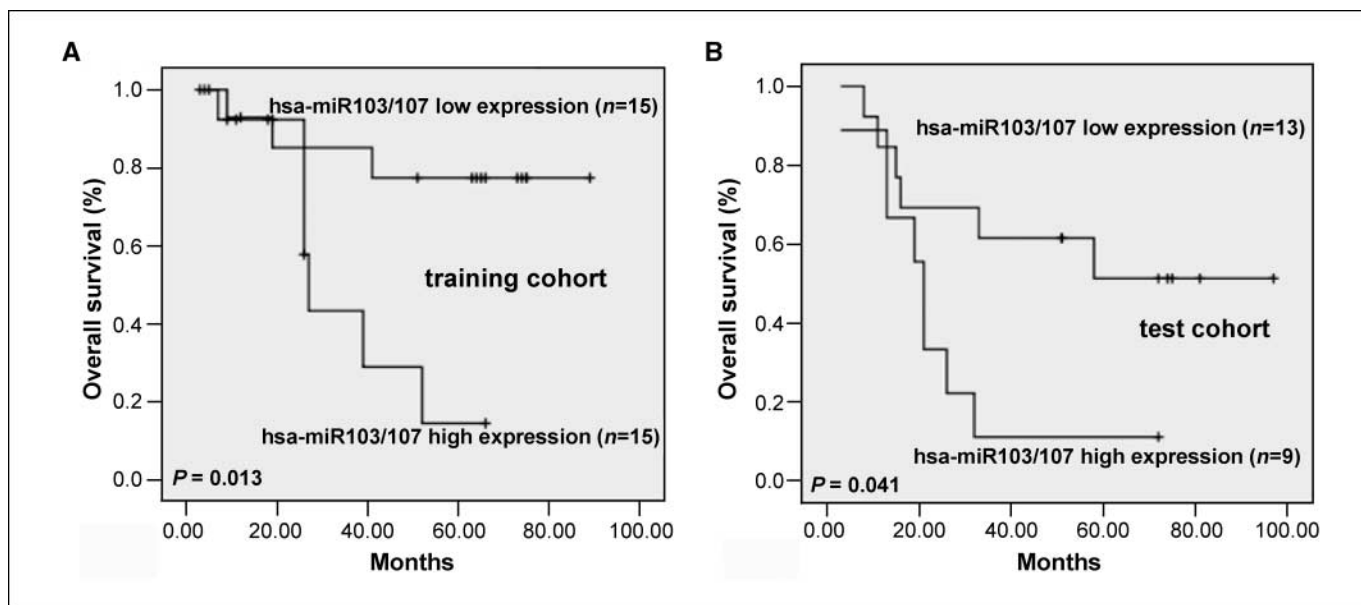

Figure 3. Kaplan-Meier overall survival curves for esophageal cancer patients. The survival curves are shown for patient samples from the training cohort (*A*) and the test cohort (*B*). Low expression of hsa-miR-103/107 correlated with a high overall survival rate and high expression correlated with a low overall survival rate.

clinicopathologic factors (age, gender, tobacco, alcohol, T, N, and TNM) revealed prognostic significance for N, TNM, and hsa-miR-103/107. These three significant variables ($P < 0.05$) were further entered into a multivariate Cox model, which indicated that both high hsa-miR-103/107 expression ($P = 0.047$) and TNM ($P = 0.002$) were strongly associated with a poor patient outcome (Table 2).

Validation of microarray data by quantitative RT-PCR analysis. Our oligonucleotide microarray-based miRNA detection platform was constructed by CapitalBio, and we undertook miRNA expression analysis according to their instructions. Several previous in depth comparative studies between microarray platforms and analysis procedures have indicated the very high reproducibility, sensitivity, and specificity of similar expression microarrays using their recommended procedures (38, 39). Further validation of the hsa-miR-103/107 expression trends were determined by quantitative RT-PCR in 11 cases. The miRNAs were found to have the same expression trends as seen by microarray analysis, with a reasonable correlation between the quantities of the transcripts measured by both microarray and quantitative RT-PCR analysis methods (Supplementary Fig. S2).

Discussion

Although a number of different microarray platforms have been developed for the quantitative assay of miRNA expression (18–23), we used a newly designed microarray platform specific for the analysis of the expression of some 509 mammalian miRNAs. The platform and assay are similar in many respect to other spotted oligonucleotide microarray designs (19) but have several important differences in application. A modified spotting buffer and an advanced hybridization system were used in this study. These measures have both previously shown large improvements in the local signal intensity and global signal uniformity as well as elimination of the doughnut spots commonly seen on spotted oligonucleotide arrays. These improvements are believed to be due to better blocking of the slide surface chemistry (31). A detailed assessment of the quality control and reproducibility of this new miRNA microarray platform has been published recently (40).

Recently, Nelson and colleagues (20) reported the first analysis of miRNA from formalin-preserved paraffin-embedded tissues by use of a RNA-primed, array-based Klenow enzyme assay, allowing for analysis from archival human tissue with known clinical and pathologic information. Here, we report the analysis of miRNA from frozen esophageal tissues that have been preserved for >5 years. It was notable that the RNA moieties (total RNA) extracted from the cryoarchive-preserved tissues stored for such long periods of time were extensively degraded (evidencing periods of nonideal preservation), whereas total RNA extracted from fresh tissues showed no such degradation. Although such nonideal tissue preservation is not lauded, obtaining intact clinical tissues from archives is a problem that has been noted previously (26). Nonetheless, our analysis of these tissues indicated that the short miRNA species were relatively stable during this cryostorage, compared with long mRNA and rRNA molecules. Further independent validation has also revealed the miRNA expression correlation coefficient (R^2 value) among fresh intact RNA and extensively degraded total RNA preparations was at least 0.924 after examining total RNA extracted from HepG2 and HEK293 cell lines and mouse liver tissue (data not shown). It is important to be able to perform accurate and informative analysis long after the patient surgery, as both survival time and other clinical data have accumulated, allowing prognostic analysis, and new resections from potentially recurrent conditions can also be compared with initially preserved tissues. The clinical information allowed us *post hoc* to analyze the potential influence of each miRNA on the disease prognosis of the cancer patient. Here, low expression of hsa-miR-103/107 was found to correlate strongly with long overall patient survival periods, and thus, these miRNAs might constitute a useful diagnostic tool or a potential drug target for esophageal cancer management.

Several other recent studies have reported the relevance of particular miRNAs to the progression of particular tumors. Calin et al. (27) reported that a unique 13-miRNA expression signature (hsa-miR-15a, hsa-miR-195, hsa-miR-221, miR-23b, miR-155, miR-223, miR-29a-2, miR-24-1, miR-29b-2, miR-146, miR-16-1, miR-16-2, and miR-29c) was a prognostic indicator of chronic lymphocytic leukemia. Yanaihara et al. (28) found that expression levels of the five miRNAs (hsa-mir-155, hsa-mir-17-3p, hsa-mir-let-7a-2, hsa-mir-145, and hsa-mir-21) were statistically altered in lung cancers, and these also had a prognostic effect on patient survival. Roldo et al. (41) showed that the expression of has-miR-103 and has-miR-107 and lack of expression of has-miR-155 could discriminate

Table 2. Postoperative survival of patients with ESCC in relation to clinicopathologic characteristics and miRNA expression analyzed by the Cox proportional hazard regression model in all 52 cases

		Univariate analysis		Multivariate analysis	
		Hazard ratio (95% CI)	P	Hazard ratio (95% CI)	P
Gender	Male/female	1.473 (0.438–4.960)	0.532		
Age	$\sim 49 \sim 67 \sim$	1.106 (0.581–2.106)	0.760		
Tobacco	Y/N	1.599 (0.634–4.033)	0.320		
Alcohol	Y/N	1.293 (0.578–2.899)	0.532		
T	$T_1/T_2/T_3$	1.917 (0.836–4.396)	0.124		
N	N_1/N_0	5.015 (2.126–11.826)	<0.001	1.235 (0.245–6.219)	0.798
TNM	I/II/III	4.663 (2.104–10.337)	<0.001	3.466 (1.555–7.728)	0.002
miRNA-103/107	High/low	3.621 (1.531–8.565)	0.003	2.604 (1.013–6.692)	0.047

Abbreviation: 95% CI, 95% confidence interval.

pancreatic tumors from normal. Here, for the first time, we identified that high expression of hsa-miR-103/107 has a negative prognostic effect on the esophageal cancer patient survival. Recently Sugito et al. (42) reported that the miRNA processing enzyme (*RNASEN*) was elevated in a proportion of ESCC and high *RNASEN* expression correlates with poor prognosis in ESCC. These findings, together with our evidence of distinctive miRNA profiles in esophageal cancers, suggest strongly that altered metabolism of particular miRNAs plays a role in esophageal cancer development.

MicroRNAs are a class of regulatory RNAs that function primarily by targeting specific mRNAs for degradation or inhibition of translation and, thus, decrease the expression of the resulting protein, and their role in tumor development would presumably be through the regulation of their target protein genes (10, 11). For some miRNAs however, their altered expression profile is not causal and might simply indicate a changed transcriptional coregulation of the miRNA and the mRNAs of cancer-related genes (43). The target genes of these causal miRNAs may be tumor suppressor genes or other genes related to oncogenes, such as growth factors, growth factor receptors, signal transducers, transcription factors, programmed cell death regulators, genes that control cell division, or genes that repair DNA. Several publications have presented algorithms with which to identify putative targets for miRNA (44–47). We used all of these algorithms to predict the putative target genes (listed in Supplementary Table S2) of the survival-related miRNAs (hsa-miR-103/107). We also conducted a bioinformatics analysis grouping the predicted targets of hsa-mir-103/107 by using the Gene Ontology, Biocarta, KEGG, and GenMAPP databases. Among the putative target genes, *YWHAH* is a tumor suppressor and regulates the cell cycle, *TGFBR3* is involved in the transforming growth factor β signaling pathway, *AXIN2* is involved in the Wnt signaling pathway, *TAF5* is a transcription factor and *CAPZA2* is involved in cell motility, and several may be involved in esophageal cancer development via different mechanisms. For example, *YWHAH* gene is known to interact with tumor suppressors and to regulate the cell cycle (48). Other reports have shown that miR-107 is up-regulated in tumors of the gastroenterological system, such as the colon, pancreas, and stomach (41, 49). The cancer tissue we investigated also belongs to the digestive system, and our results also indicate that high expression of miR-103/miR-107 in esophageal cancer can be associated with a poor prognosis. In contrast, Garzon and colleagues (50) reported that the up-regulation of miR-107 could induce promyelocytic differentiation, suggesting that miR-107 is a "protective" miRNA in these promyelocytic cells. Importantly, these observations for different gastroenterological tissues and promyelocytic tumors suggest that the functions of miR-103/107 need to be further explored in particular tumors. The current report may also provide impetus toward identifying more genes associated with ESCC.

In summary, we have investigated the miRNA expression profile of esophageal cancers with cryopreserved archival tissues stored for the periods of 5 years or more. Our microarray analyses revealed that 46 miRNAs are differently expressed between the cancerous and adjacent normal tissues and that a minimal set of 7 of them can distinguish malignant from normal esophageal tissues. Some miRNAs showed correlation with several different clinicopathologic classifications. Here, hsa-miR-103/107 showed a strong correlation between low expression levels and a high overall and disease-free survival periods for esophageal cancer patients by univariate analysis, as well as by multivariate analysis. These results should provide impetus to examine new molecular mechanisms that may lead to development of esophageal cancers, and hsa-miR-103/107 might prove useful for the diagnostic analysis of esophageal cancers using a simple, fast, and economical approach such as RT-PCR.

Acknowledgments

References

1. Enzinger PC, Mayer RJ. Esophageal cancer. N Engl J Med 2003;349:2241–52.
2. Hollstein MC, Metcalf RA, Welsh JA, Montesano R, Harris CC. Frequent mutation of the p53 gene in human esophageal cancer. Proc Natl Acad Sci U S A 1990;87:9958–61.
3. Jiang W, Zhang YJ, Kahn SM, et al. Altered expression of the cyclin D1 and retinoblastoma genes in human esophageal cancer. Proc Natl Acad Sci U S A 1993;90:9026–30.
4. Hu N, Wang C, Han XY, et al. Evaluation of BRCA2 in the genetic susceptibility of familial esophageal cancer. Oncogene 2004;23:852–8.
5. Gratas C, Tohma Y, Barnas C, Taniere P, Hainaut P, Ohgaki H. Up-regulation of Fas (APO-1/CD95) ligand and down-regulation of Fas expression in human esophageal cancer. Cancer Res 1998;58:2057–62.
6. Shih CH, Ozawa S, Ando N, Ueda M, Kitajima M. Vascular endothelial growth factor expression predicts outcome and lymph node metastasis in squamous cell carcinoma of the esophagus. Clin Cancer Res 2000;6:1161–8.
7. Mizushima T, Nakagawa H, Kamberov YG, Wilder EL, Klein PS, Rustgi AK. Wnt-1 but not epidermal growth factor induces β-catenin/T-cell factor-dependent transcription in esophageal cancer cells. Cancer Res 2002;62:277–82.
8. Yang L, Leung AC, Ko JM, et al. Tumor suppressive role of a 2.4 Mb 9q33–34 critical region and DEC1 in esophageal squamous cell carcinoma. Oncogene 2005;24:697–705.
9. Bartel DP. MicroRNAs: genomics, biogenesis, mechanism, and function. Cell 2004;116:281–97.
10. Esquela-Kerscher A, Slack FJ. Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer 2006;6:259–69.
11. Calin GA, Croce CM. MicroRNA signatures in human cancers. Nat Rev Cancer 2006;6:857–66.
12. Calin GA, Dumitru CD, Shimizu M, et al. Frequent deletions and down-regulation of micro-RNA genes miR15 and miR16 at 13q14 in chronic lymphocytic leukemia. Proc Natl Acad Sci U S A 2002;99:15524–9.
13. Takamizawa J, Konishi H, Yanagisawa K, et al. Reduced expression of the let-7 microRNAs in human lung cancers in association with shortened postoperative survival. Cancer Res 2004;64:3753–6.
14. Murakami Y, Yasuda T, Saigo K, et al. Comprehensive analysis of microRNA expression patterns in hepatocellular carcinoma and non-tumorous tissues. Oncogene 2006;25:2537–45.
15. Calin GA, Sevignani C, Dumitru CD, et al. Human microRNA genes are frequently located at fragile sites and genomic regions involved in cancers. Proc Natl Acad Sci U S A 2004;101:2999–3004.
16. Saito Y, Liang G, Egger G, et al. Specific activation of microRNA-127 with downregulation of the proto-oncogene BCL6 by chromatin-modifying drugs in human cancer cells. Cancer Cell 2006;9:435–43.
17. Voorhoeve PM, le Sage C, Schrier M, et al. A genetic screen implicates miRNA-372 and miRNA-373 as oncogenes in testicular germ cell tumors. Cell 2006;124:1169–81.
18. Liu CG, Calin GA, Meloon B, et al. An oligonucleotide microchip for genome-wide microRNA profiling in human and mouse tissues. Proc Natl Acad Sci U S A 2004;101:9740–4.
19. Thomson JM, Parker J, Perou CM, Hammond SM. A custom microarray platform for analysis of microRNA gene expression. Nat Methods 2004;1:47–53.
20. Nelson PT, Baldwin DA, Scearce LM, Oberholtzer JC, Tobias JW, Mourelatos Z. Microarray-based, high-throughput gene expression profiling of microRNAs. Nat Methods 2004;1:155–61.
21. Barad O, Meiri E, Avniel A, et al. MicroRNA expression detected by oligonucleotide microarrays: system establishment and expression profiling in human tissues. Genome Res 2004;14:2486–94.
22. Babak T, Zhang W, Morris Q, Blencowe BJ, Hughes TR. Probing microRNAs with microarrays: tissue specificity and functional inference. RNA 2004;10:1813–9.
23. Shingara J, Keiger K, Shelton J, et al. An optimized isolation and labeling platform for accurate microRNA expression profiling. RNA 2005;11:1461–70.

24. Lu J, Getz G, Miska EA, et al. MicroRNA expression profiles classify human cancers. Nature 2005;435:834–8.
25. Chen C, Ridzon DA, Broomer AJ, et al. Real-time quantification of microRNAs by stem-loop RT-PCR. Nucleic Acids Res 2005;33:e179.
26. Haller AC, Kanakapalli D, Walter R, Alhasan S, Eliason JF, Everson RB. Transcriptional profiling of degraded RNA in cryopreserved and fixed tissue samples obtained at autopsy. BMC Clin Pathol 2006;6:9.
27. Calin GA, Ferracin M, Cimmino A, et al. A MicroRNA signature associated with prognosis and progression in chronic lymphocytic leukemia. N Engl J Med 2005;353: 1793–801.
28. Yanaihara N, Caplen N, Bowman E, et al. Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 2006;9:189–98.
29. Xie X, Lu J, Kulbokas EJ, et al. Systematic discovery of regulatory motifs in human promoters and 3′ UTRs by comparison of several mammals. Nature 2005;434: 338–45.
30. Watanabe T, Takeda A, Mise K, et al. Stage-specific expression of microRNAs during Xenopus development. FEBS Lett 2005;579:318–24.
31. Patterson TA, Lobenhofer EK, Fulmer-Smentek SB, et al. Performance comparison of one-color and two-color platforms within the MicroArray Quality Control (MAQC) project. Nat Biotechnol 2006;24:1140–50.
32. Eisen MB, Spellman PT, Brown PO, Botstein D. Cluster analysis and display of genome-wide expression patterns. Proc Natl Acad Sci U S A 1998;95:14863–8.

33. Tusher VG, Tibshirani R, Chu G. Significance analysis of microarrays applied to the ionizing radiation response. Proc Natl Acad Sci U S A 2001;98:5116–21.
34. Dai JJ, Lieu L, Rocke D. Dimension reduction for classification with gene expression microarray data. Stat Appl Genet Mol Biol 2006;5:Article6.
35. Wu B, Abbott T, Fishman D, et al. Comparison of statistical methods for classification of ovarian cancer using mass spectrometry data. Bioinformatics 2003;19: 1636–43.
36. Efron B, Gong G. A leisurely look at the bootstrap, the jackknife, and cross-validation. Am Stat 1983;37:36–48.
37. Blot WJ. Esophageal cancer trends and risk factors. Semin Oncol 1994;21:403–10.
38. Shi L, Reid LH, Jones WD, et al. The MicroArray Quality Control (MAQC) project shows inter- and intra-platform reproducibility of gene expression measurements. Nat Biotechnol 2006;24:1151–61.
39. Guo Y, Guo HY, Zhang L, et al. Genomic analysis of anti-hepatitis B virus (HBV) activity by small interfering RNA and lamivudine in stable HBV-producing cells. J Virol 2005;79:14392–403.
40. Luo MY, Tian ZG, Xu Z, et al. Construction and application of a microarray for profiling microRNA expression. Prog Biochem Biophys 2007;34:31–41.
41. Roldo C, Missiaglia E, Hagan JP, et al. MicroRNA expression abnormalities in pancreatic endocrine and acinar tumors are associated with distinctive pathologic features and clinical behavior. J Clin Oncol 2006;24: 4677–84.

42. Sugito N, Ishiguro H, Kuwabara Y, et al. RNASEN regulates cell proliferation and affects survival in esophageal cancer patients. Clin Cancer Res 2006;12: 7322–8.
43. Hornstein E, Shomron N. Canalization of development by microRNAs. Nat Genet 2006;38:S20–4.
44. Lewis BP, Shih IH, Jones-Rhoades MW, Bartel DP, Burge CB. Prediction of mammalian microRNA targets. Cell 2003;115:787–98.
45. John B, Enright AJ, Aravin A, Tuschl T, Sander C, Marks DS. Human microRNA targets. PLoS Biol 2004;2: e363.
46. Krek A, Grun D, Poy MN, et al. Combinatorial microRNA target predictions. Nat Genet 2005;37:495–500.
47. Griffiths-Jones S, Grocock RJ, van Dongen S, Bateman A, Enright AJ. miRBase: microRNA sequences, targets and gene nomenclature. Nucleic Acids Res 2006;34: D140–4.
48. Yu T, Robb VA, Singh V, Gutmann DH, Newsham IF. The 4.1/ezrin/radixin/moesin domain of the DAL-1/ Protein 4.1B tumour suppressor interacts with 14–3-3 proteins. Biochem J 2002;365:783–9.
49. Volinia S, Calin GA, Liu CG, et al. A microRNA expression signature of human solid tumors defines cancer gene targets. Proc Natl Acad Sci U S A 2006;103: 2257–61.
50. Garzon R, Pichiorri F, Palumbo T, et al. MicroRNA gene expression during retinoic acid-induced differentiation of human acute promyelocytic leukemia. Oncogene 2007;26:4148–57.