Published Ahead of Print on March 1, 2010 as 10.1200/JCO.2009.24.9342
The latest version is at http://jco.ascopubs.org/cgi/doi/10.1200/JCO.2009.24.9342

JOURNAL OF CLINICAL ONCOLOGY ORIGINAL REPORT

Serum MicroRNA Signatures Identified in a Genome-Wide Serum MicroRNA Expression Profiling Predict Survival of Non–Small-Cell Lung Cancer

Zhibin Hu, Xi Chen, Yang Zhao, Tian Tian, Guangfu Jin, Yongqian Shu, Yijiang Chen, Lin Xu, Ke Zen, Chenyu Zhang, and Hongbing Shen

From the Department of Epidemiology and Biostatistics, Cancer Center; and Departments of Oncology and Thoracic Surgery, First Affiliated Hospital of Nanjing Medical University, Nanjing Medical University; State Key Laboratory of Pharmaceutical Biotechnology, School of Life Sciences, Nanjing University; and the Department of Thoracic Surgery, Jiangsu Cancer Hospital, Nanjing, China.

Submitted July 5, 2009; accepted November 24, 2009; published online ahead of print at www.jco.org on March 1, 2010.

Supported in part by the National Natural Science Foundation Grants No. 30570731 and 30730080 and by the National Basic Research Program No. 973 Grants No. 2006CB503909, 2004CB518603, and 2002CB512902.

Z.H. and X.C. contributed equally to this work.

Authors' disclosures of potential conflicts of interest and author contributions are found at the end of this article.

Corresponding authors: Hongbing Shen, MD, PhD, Department of Epidemiology and Biostatistics, Cancer Center, Nanjing Medical University, 140 Hanzhong Rd, Nanjing 210029, China; e-mail: hbshen@njmu.edu.cn.

0732-183X/10/2899-1/$20.00

DOI: 10.1200/JCO.2009.24.9342

ABSTRACT

Purpose
Recent findings that human serum contains stably expressed microRNA (miRNA) have revealed a great potential of serum miRNA signature as disease fingerprints to predict survival. We used genome-wide serum miRNA expression analysis to investigate the role of serum miRNA in predicting prognosis of non–small-cell lung cancer (NSCLC).

Patients and Methods
To control disease heterogeneity, we used patients with stages I to IIIa lung adenocarcinoma and squamous cell carcinoma, who were treated with both operation and adjuvant chemotherapies. In the discovery stage, Solexa sequencing followed by individual quantitative reverse transcriptase polymerase chain reaction (qRT-PCR) assays was used to test the difference in levels of serum miRNAs between 30 patients with longer survival (alive and mean survival time, 49.54 months) and 30 patients with shorter survival matched by age, sex, and stage (dead and mean survival time, 9.54 months). The detected serum miRNAs then were validated in 243 patients (randomly classified into two subgroups: n = 120 for the training set, and n = 123 for the testing set).

Results
Eleven serum miRNAs were found to be altered more than five-fold by Solexa sequencing between longer-survival and shorter-survival groups, and levels of four miRNAs (ie, miR-486, miR-30d, miR-1 and miR-499) were significantly associated with overall survival. The four-miRNA signature also was consistently an independent predictor of overall survival for both training and testing samples.

Conclusion
The four-miRNA signature from the serum may serve as a noninvasive predictor for the overall survival of NSCLC.

J Clin Oncol 28. © 2010 by American Society of Clinical Oncology

INTRODUCTION

Lung cancer is the leading cause of cancer deaths worldwide, and approximately 80% of the patients have non–small-cell lung cancer (NSCLC). Because lung cancer is a deadly disease that has a 5-year survival rate of less than 15%, prognostic assessment of the patient is essential for the choice of better therapeutic strategies. One of the major clinical determinants in NSCLC prognosis is tumor extension, roughly characterized by the stage. However, a large variability in disease outcome has been observed for a subset of patients with the same clinical features, and the current staging systems are inadequate to predict the treatment outcome of NSCLC.

It has been suggested that discovery and application of molecular biomarkers that incorporate with traditional cancer staging and prognosis signature could improve the management of patients with NSCLC.[1] Advances in genomics and proteomics have generated many new candidate markers with potential clinical values, particularly the gene-expression profiling by microarray or real-time reverse transcriptase polymerase chain reaction (RT-PCR).[2-4] Recently, the discovery of microRNAs (miRNAs) has opened new avenues for cancer diagnosis and prediction of treatment response.[5-6]

miRNAs are an abundant class of small, non–protein-coding RNAs that function as negative gene regulators.[7-8] Because a single miRNA can regulate hundreds of downstream genes with different biologic entities, the information gained from miRNA profiling may provide more accurate classification of cancer subtypes than the use of expression profiles of protein-coding genes.[6,9] In lung cancer, it has been shown that miRNA expression profiles and

specific miRNAs in lung tissue are correlated with disease prognosis and clinical outcome.[10,11]

Recently, we discovered with excitement that human serum/plasma contains a large amount of stable miRNAs and that unique expression patterns of serum miRNAs under a specific disease condition imply a great potential of serum miRNA profiling as the fingerprint for diseases.[12] Consistently, other studies also have suggested that serum miRNAs are novel noninvasive biomarkers for cancers.[13-15] In this study, we hypothesize that there is a serum miRNA profile that can be used as a fingerprint to predict clinical outcome of NSCLC. To address this hypothesis, we screened serum miRNAs by using Solexa sequencing followed by an extensively self-validated study that uses individual quantitative RT-PCR (qRT-PCR) assays in a cohort of 303 patients with stages I to IIIa NSCLC.

PATIENTS AND METHODS

Detailed patients recruitment and methodologic issues are described in the Appendix (online only). Patient characteristics and clinical features were partly described previously.[16] To control disease heterogeneity, only patients with stages I to IIIa lung cancer with adenocarcinoma and squamous cell carcinoma treated with both surgical operation and adjuvant chemotherapies were included on this study. Overall, 303 patients who met all criteria and had completed follow-up and clinical information were included in this analysis. Among them, 60 patients were selected for the discovery stage for Solexa sequencing of serum miRNAs, 30 patients who survived more than 30 months (average, 49.54 months; range, 34.6 to 61.8 months) on the last follow-up were classified as the longer-survival group, whereas 30 patients who had survival times less than 25 months (average, 9.54 months; range, 2.0 to 22.5 months) were classified as the shorter-survival group. The remaining 243 patients with NSCLC were randomly classified as either training (n = 120) or testing (n = 123) sets by using computer-generated random numbers.

RNA isolation, Solexa sequencing, and qRT-PCR assay were described previously.[12] In each step from serum purification to qRT-PCR, an equal volume of serum was processed. Because there was no consensus on the use of housekeeping miRNA for serum qRT-PCR analysis, miRNA expression levels all were controlled with a healthy donor sample. For reverse transcription, the same control sample was processed together with the testing samples according to their plate formats. In each 96-well plate, the control sample was assayed in sextuple in parallel with other patient samples, and the ratios of serum miRNAs were calculated by using the equation $2^{-\triangle CT}$, in which $^{\triangle CT} = \text{cycle threshold (CT)}_{\text{patient}} - \text{CT}_{\text{control}}$. In this study, the coefficient of variation (CV) of CT values for the control sample between different plates for different miRNAs was small and comparable (CV = 0.021 for miR-486, 0.061 for miR-30d; CV = 0.047 for miR-1 and 0.041 for miR-499), which indicated that the reverse transcription and quantitative PCR were comparably efficient in all reactions.

The associations between overall survival and serum miRNA expression levels were estimated by using the Kaplan-Meier method, log-rank test, and Cox proportional hazard regression models. We assigned each patient a risk score according to a linear combination of the expression level of the miRNA, weighted by the regression coefficient from the training samples.[2,11,17] The risk score was calculated as follows: risk-score = (0.969 × expression level of miR-486) + (0.973 × expression level of miR-30d) + (−0.650 × expression level of miR-1) + (−0.815 × expression level of miR-499). The Cox stepwise regression model and stratification analyses also were conducted. All the statistical analyses were performed with Statistical Analysis System software (version 9.1.3; SAS Institute, Cary, NC).

RESULTS

Detection of Serum miRNAs and Their Associations With Survival of Patients With NSCLC

As shown in Appendix Table A2 (online only), the 30 patients in the longer-survival group and the 30 patients in the shorter-survival group in the discovery stage were exactly matched on stage and smoking status, and the frequencies of histologic type, age, and sex were not significantly different between the two groups. Overall, 109 miRNAs were detected by Solexa sequencing from the longer-survival group, whereas 101 were detected from the shorter-survival group. We selected miRNAs that satisfied two criteria for additional qRT-PCR validation: having at least 50 copies in either longer-survival or shorter-survival groups; and showing five-fold altered expression between the two pooled samples. As a result, 11 serum miRNAs were identified and subjected to additional analyses (Table 1). In addition,

Table 1. Copies of Solexa Sequencing and qRT-PCR Expression of the 11 miRNAs

	Long Survival Group					Short Survival Group					Overall	
miRNA	No.	Copies*	Mean	SD	Median	No.	Copies*	Mean	SD	Median	Ratio†	*P*‡
miR-486	30	163	0.658	0.361	0.669	30	3205	1.129	0.542	1.065	0.051	< .001
miR-22	30	18	0.722	1.225	0.375	30	209	1.103	2.401	0.254	0.086	.443
miR-30d	30	28	0.455	0.371	0.341	30	165	1.353	0.738	1.122	0.170	< .001
miR-21	30	230	1.490	0.544	1.480	30	44	1.239	0.501	1.239	5.227	.069
miR-26b	30	56	1.423	1.337	0.943	30	10	1.240	0.973	0.793	5.600	.545
let-7i	30	153	2.249	1.805	1.789	30	26	1.845	1.139	1.563	5.885	.304
miR-378	30	212	2.946	5.963	0.740	30	34	4.407	11.335	1.042	6.235	.535
miR-1	30	2918	2.918	3.818	1.182	30	407	0.408	0.608	0.180	7.170	.001
miR-206	30	1979	1.471	0.898	1.450	30	196	1.500	1.023	1.379	10.097	.907
miR-146b	30	81	1.309	1.012	1.114	30	6	1.299	0.823	1.315	13.500	.966
miR-499	30	112	1.096	0.465	1.039	30	6	0.683	0.269	0.712	18.667	< .001
let-7a	30	1536	0.857	0.384	0.776	30	742	0.907	0.380	1.035	2.070	.615
let-7g	30	176	1.430	0.733	1.382	30	59	1.410	0.775	1.257	2.983	.919

Abbreviations: qRT-PCR, qualitative reverse transcriptase polymerase chain reaction; miRNA, microRNA; SD, standard deviation.
*Solexa sequencing results of 30 pooled samples.
†Copies of long survival group ÷ copies of short survival group.
‡Student *t* test.

Table 2. Serum Expression Levels of Four miRNAs and Survival of Patients With NSCLC in All Sample Sets

miRNA	Patients (N = 303)	Deaths (n = 113)	MST (months)	Log-Rank *P*	Adjusted HR	95% CI	*P*
miR-486							
Low, ≤ 0.795	149	44	Not reached		1.00		
High, > 0.795	154	69	31.30	.0002	2.01	1.36 to 2.98	< .001
miR-30d							
Low, ≤ 0.680	147	35	Not reached		1.00		
High, >0.680	156	78	22.37	< .0001	4.26	2.81 to 6.47	< .001
miR-1							
High, > 0.675	149	31	Not reached		1.00		
Low, ≤ 0.675	154	82	22.87	< .0001	3.60	2.36 to 5.47	< .001
miR-499							
High, > 0.750	166	45	Not reached		1.00		
Low, ≤ 0.750	137	68	23.73	< .0001	2.44	1.65 to 3.60	< .001
No. of high-risk miRNA							
0-1	100	14	Not reached		1.00		
2	108	29	50.67		3.14	1.65 to 5.97	
3	74	50	17.83		16.52	8.62 to 31.68	
4	21	20	12.27	< .0001	34.13	16.28 to 71.56	< .001*

NOTE. Analysis was adjusted for age, sex, smoking status, histology, and stage.
Abbreviations: miRNA, microRNA; NSCLC, non–small-cell lung cancer; MST, median survival time.
**P* for trend.

two let-7 family members (ie, let-7a and let-7g) with two-fold expression changes between the two groups were included also, because their expressions in tissue specimen were reported to be correlated with lung cancer survival.[18,19] We then performed individual qRT-PCR assay to quantify each of the identified miRNAs. As shown in Table 1, the levels of four miRNAs (ie, miR-486, miR-30d, miR-1 and miR-499) were significantly different between the longer-survival and shorter-survival groups ($P < .001$ for miR-486; $< .001$ for miR-30d; .001 for miR-1; and $< .001$ for miR-499).

Hazard ratios (HRs) from the Cox regression analysis showed that the expression levels of the four miRNAs were significantly associated with cancer death in a dose-dependent manner (for trend = .001 for miR-486, < .001 for miR-30d, .003 for miR-1, and < .001 for miR-499). The difference in the median survival time (MST) of lung cancer between different expression levels of the four miRNAs in quartiles were all statistically significant (log-rank test $P = .009$ for miR-486, $< .001$ for miR-30d, $< .001$ for miR-1, and .001 for miR-499; Appendix Table A3, online only).

Association Between the Expression of the Four miRNAs in Serum and Patient Survival

We additionally tested the predictive effects of the four miRNAs on lung cancer survival among 243 patients with NSCLC, randomly grouped into one training data set and one testing data set (Appendix Table A2). From the training set, high serum expression levels of miR-486 (median, 0.795) and miR-30d (median, 0.680), and low serum expression levels of miR-1 (median, 0.675), miR-499 (median, 0.750) were all individually associated with unfavorable survival (Appendix Table A4 and Fig A1A, online only). When these threshold values were applied to the testing set, comparable log-rank *P* values and HRs, respectively, also were observed (Appendix Table A4 and Fig A1B). We then took the four miRNAs into simultaneous consideration and found that patients carrying two or more high-risk miRNAs had significantly increased probability of shortened survival than those carrying zero or one high-risk miRNA in both the training and testing sets (Appendix Table A4 and Figs A1A, A1B).

Results for all the 303 patient samples, including the discovery data set, also were similar to those from the training set. As listed in Table 2, patients with high serum expression levels of risk miRNAs and low serum expression levels of protective miRNAs had significantly lower MSTs (31.30 months for miR-486, 22.37 months for miR-30d, 22.87 months for miR-1, and 23.73 months for miR-499; all *v* not reached; Appendix Fig A1C). In the combined analysis, we found that patients carrying two or more high-risk miRNAs had significantly increased probability of cancer death in a dose-dependent manner (log-rank test $< .0001$; HR = 3.14; 95% CI, 1.65 to 5.97 for two high-risk miRNAs carriers; HR, 16.52; 95% CI, 8.62 to 31.68 for 3 high-risk miRNAs carriers; and HR, 34.13; 95% CI, 16.28 to 71.56 for four high-risk miRNAs carriers; Table 2) compared with those carrying zero or one high-risk miRNA.

Serum Four-miRNA Signature and Patient Survival

We then used the four miRNAs to construct a signature by a risk-score method. The patients in the training set were ranked according to their risk scores and divided into high-risk or low-risk groups by using the median risk score as the cutoff point (Table 3). Patients with a high-risk four-miRNA signature had significantly shortened MST (20.30 months *v* not reached; $P < .001$) and a 10.74-fold increased HR (95% CI, 4.28 to 26.96) for cancer death. The same risk-score formula and cutoff point obtained from the training set were directly applied to the 123 patients in the testing set and to the all-combined 303 patients. Similar to the findings from the training set, patients with a high-risk miRNA signature had significantly shortened MST than patients with a low-risk miRNA signature (23.97 *v* 43.17 months; $P < .001$ for the testing set; and 18.43 months *v* not reached; $P < .001$ for the all-combined 303 patients) and elevated HR (HR, 6.35; 95% CI, 3.07 to 13.15 for the testing set; and HR, 9.31; 95% CI, 5.78 to 14.98 for the all-combined 303 patients).

Table 3. Risk Score and NSCLC Survival for Training, Testing, and All Combined Data Sets

miRNA by Data Set	Patients	Deaths	MST (months)	Log-Rank *P*	Adjusted HR	95% CI
Training set						
No. of patients	120	40				
Low, ≤ 0.415	60	7	Not reached		1.00	
High, > 0.415	60	33	20.3	< .001	10.74	4.28 to 26.96
Testing set						
No. of patients	123	43				
Low, ≤ 0.415	75	16	43.17		1.00	
High, > 0.415	48	27	23.97	< .001	6.35	3.07 to 13.15
All data sets						
No. of patients	303	113				
Low, ≤ 0.415	167	26	Not reached		1.00	
High, > 0.415	136	87	18.43	< .001	9.31	5.78 to 14.98

NOTE. Analysis was adjusted for age, sex, smoking status, histology, and stage.
Abbreviations: NSCLC, non–small-cell lung cancer; miRNA, microRNA; MST, median survival time.

To test the stability of the cutoff point, we also performed 1,000 times of cross-validations. Each time, we generated a risk-score formula on 120 randomly selected participants from the training and testing sets. The average and standard deviation of the 1,000 threshold values were 0.455 ± 0.578. On the basis of these cutoff points, the all-combined 303 patients were classified into high-risk or low-risk groups. The mean with an empirical standard error of the 1,000 HRs was 8.29 ± 1.84, and the empiric 95% CI was between 4.48 and 10.59.

We showed the distribution of serum miRNA expression levels, patient risk scores, and survival time of the all-combined 303 patients in Fig 1. Patients with high-risk scores tended to have high-risk miRNAs, whereas patients with low-risk scores tended to have more protective miRNAs. Patients with high-risk scores died sooner than those with low-risk scores.

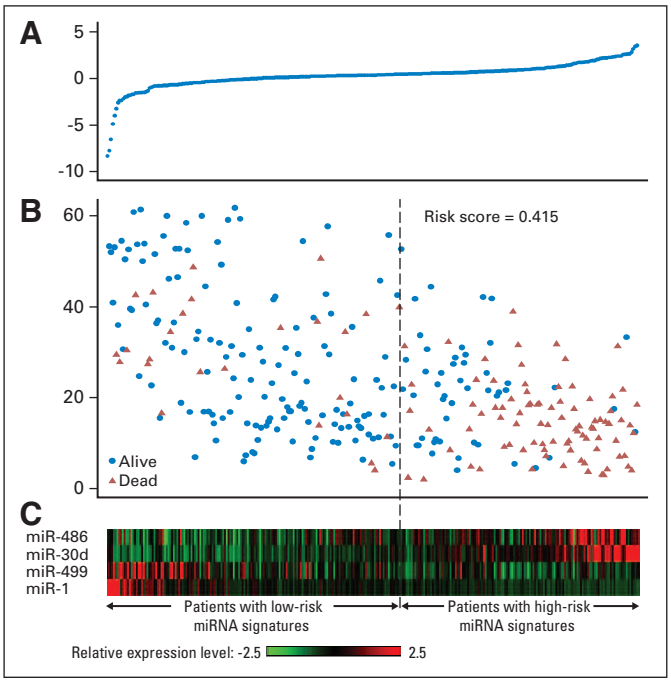

Fig 1. Risk score analysis of 303 early-stage patients with non–small-cell lung cancer (NSCLC). (A) risk-score distribution; (B) patients' survival status and time; (C) color-gram of miRNA expression profiles of patients with NSCLC, rows represent miRNAs, and columns represent patients.

Serum miRNA Signatures and Patient Survival Independent From Stage and Histology

To investigate whether the miRNA signatures could distinguish high-risk versus low-risk groups of patients by stage (ie, NSCLC stage I, II, or IIIa) and histology subtypes (ie, adenocarcinoma or squamous cell carcinoma), we performed stratified analyses. We found that the increased risk of cancer death associated with the high-risk four-miRNA signature was comparable among patients with stage I (HR, 12.79; 95% CI, 4.67 to 35.03), II (HR, 9.38; 95% CI, 3.00 to 29.36), and IIIa (HR, 8.99; 95% CI, 4.75 to 17.03) carcinomas and between patients with adenocarcinoma (HR, 9.34; 95% CI, 5.12 to 17.04) and squamous cell carcinoma (HR, 12.11; 95% CI, 4.86 to 30.17). We also performed stepwise Cox proportional hazards analyses with selected demographic characteristics, clinical features, and the four miRNAs or the four-miRNA signature. The four miRNAs or the four-miRNA signature, together with stage remained in the regression model with a significance level of .05 for entering and .10 for removing of the variables (*P* all < .001), suggesting that either four miRNAs individually or the four-miRNA signature as a group may be used as biomarkers to predict the lung cancer survival in this study population (Appendix Table A5, online only).

DISCUSSION

NSCLC is a heterogeneous disease. The current standard of treatment for patients with stage Ia NSCLC is surgical resection, but a subgroup of these patients might benefit from adjuvant chemotherapy; in contrast, patients with clinical stages Ib to IIIa NSCLC receive both operation and adjuvant chemotherapy, of whom some may receive potentially toxic chemotherapy unnecessarily.[20-22] Staging systems for lung cancer may have reached their limit of usefulness for predicting outcomes, and molecular biomarkers add values.[2,4,23-27] Thus, the ability to identify subgroups of patients more accurately, may refine the prognostic model and lead to more personalized cancer treatment.

In this study of lung cancer survival in Chinese, we found that the four-miRNA signature (ie, miR-486, miR-30d, miR-1 and miR-499) derived from a genome-wide serum miRNA expression profiling were significantly associated with overall survival of lung cancer, which may serve as a noninvasive predictor for prognosis of NSCLC. These serum markers, if validated in other populations and in prospective studies, will provide the opportunity for a randomized clinical trial to evaluate the benefit of identifying subgroups of patients with NSCLC who are at high risk of unfavorable survival. In addition, because of noninvasive characteristics of these serum markers and easy detection in clinical settings, such kind of miRNA signatures may be used to additionally investigate drug response and chemosensitivity of NSCLC by observing their dynamic expression levels and profiles before and after treatment.

Recent studies by both us and others implied that the unique patterns of serum miRNAs may serve as noninvasive biomarkers for cancer development and prognosis.[12-15,28] Therefore, determination of miRNA signatures in patients by using qRT-PCR to assess the profile of a small number of miRNAs as shown in this study is a clinically applicable procedure. Endogenous circulating miRNAs are stable, are well protected from RNases, and remain stable even after being subjected to harsh conditions.[12,15] Because of the simplicity and reproducibility of getting a blood sample, easily testable biomarkers found in blood serum may have a great potential in surveillance of cancer outcomes.

The source of serum miRNAs and, more importantly, the mechanisms that control the generation of serum miRNAs remain enigmatic. We have detected almost identical miRNA profilings in serum and blood cells from healthy individuals.[12] However, under the disease condition like cancer, miRNAs entered circulation dominantly from tumor cells.[12,15] The release mechanisms may be complicated and may involve tumor-derived microvesicles/exosomes.[29,30] Microvesicles are 30 to 100 nm in diameter and are shed from many different cell types under both normal and pathologic conditions.[31] They can transfer some of their contents (ie, proteins, mRNAs, and miRNAs) to other cell types and generate similar functional significance.[29,30,32,33] It was shown that some of the miRNAs were expressed at higher levels in microvesicles shed from tumor cells than inside the cells, which implies that these miRNAs may be selectively enriched in microvesicles.[30] For lung cancer, the let-7 miRNA family was the only one in tissue specimen that was consistently reported to be associated with clinical outcome.[10,11,19] However, it is not the case in the serum we tested in this study. Several suggested predictors (ie, miRNAs) from tissue specimen, like miR-20, miR-106a, miR-17-3p, miR-155, miR-145, miR-93,[10] miR-137, miR-372, and miR-182-3p[11] were not detectable by Solexa sequencing in the serum. Therefore, the predictive role of serum miRNAs could be independent from tissue specimen. In this study, clear and significant associations between noninvasive miRNA signatures from serum and lung cancer prognosis are potentially useful in clinical practice, although the mechanisms need to be additional explored.

Our understanding of miRNA expression patterns as potential biomarkers for diagnosis, prognosis, personalized therapy, and disease management is still in its infancy. A full understanding of the mRNA targets and the molecular mechanisms by which the miRNAs regulate NSCLC survival may lead to clinical application of serum miRNAs. Among the four miRNAs, only miR-1 was extensively evaluated in the context of lung cancer in the literature. MiR-1 was reported to be significantly reduced in primary lung cancer tissue and cell lines because of both genetic and epigenetic alterations.[34] Ectopic expression of miR-1 significantly reduced A549 lung cancer cell proliferation, cell migration and motility, and tumor growth in nude mice, whereas depletion of miR-1 facilitated lung cancer cell (N417) growth.[34] The downregulation of MET oncogene could be a possible mechanism by which miR-1 regulates growth and metastatic potential of these cells.[34,35] Furthermore, trichostatin A, a histone deacetylase (HDAC) inhibitor, can activate the repressed miR-1 in lung cancer cells, whereas HDACs have been identified as therapeutic targets in a variety of cancers, including lung cancer,[36] and HDAC4 is a validated target of miR-1.[34,35,37] The potential therapeutic role of miR-1 also underscored that ectopic expression of miR-1 can sensitize A549 cells to the anticancer drug doxorubicin through facilitating activation of Caspase 3, Caspase 7, and PARP-1 as well as depletion of Mcl-1.[34] As for miR-486, an increase in its expression was reported in bronchioalveolar stem cells compared with control cells,[38] whereas bronchioalveolar stem cells were suggested to be the progenitor cells of lung cancer stem cells.[39] Lung cancer stem cells are believed to be the most important players in tumor development, metastasis, therapy resistance and cancer recurrence.[39,40] Additional investigation of the regulatory mechanism of these miRNAs and their target mRNAs may improve our understanding of the molecular pathogenesis of NSCLC as well as our effectiveness in finding out potential therapy targets for NSCLC.

Our study was beneficiated from an initial genome-wide Solexa survey followed by three sets of individual qRT-PCR evaluations. The predictive effects between the relative homogeneity training and testing data sets were consistent for both individual markers and different combinations, which suggests a good internal validation (splitting samples[41]). For the discovery data set, the performance of the four miRNAs was also comparative and served as a good external validation. We also performed a bootstrap-like cross-validation on the reliability of the risk-score procedure. These results clearly indicate that, although one particular miRNA in serum may help distinguish different subgroups, a combination of a panel of miRNAs has a great potential to offer much more sensitive and specific diagnostic tests. In different lung cancer studies of protein-coding gene expression profiles, the genes associated with patient survival can vary substantially from study to study, and few genes were consistently reported.[2,4,23-27] Therefore, the generalization of the serum miRNA signatures identified in this study warrants additional studies in different ethnic populations, preferably in prospective studies for cancer outcome surveillance as well as therapy response evaluations.

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Zhibin Hu, Xi Chen, Ke Zen, Chenyu Zhang, Hongbing Shen
Financial support: Chenyu Zhang, Hongbing Shen
Administrative support: Chenyu Zhang, Hongbing Shen
Provision of study materials or patients: Yongqian Shu, Yijiang Chen, Lin Xu, Chenyu Zhang, Hongbing Shen

Collection and assembly of data: Zhibin Hu, Xi Chen, Tian Tian, Guangfu Jin, Chenyu Zhang, Hongbing Shen
Data analysis and interpretation: Zhibin Hu, Xi Chen, Yang Zhao, Ke Zen, Chenyu Zhang, Hongbing Shen
Manuscript writing: Zhibin Hu, Xi Chen, Chenyu Zhang, Hongbing Shen
Final approval of manuscript: Zhibin Hu, Xi Chen, Yang Zhao, Tian Tian, Guangfu Jin, Yongqian Shu, Yijiang Chen, Lin Xu, Ke Zen, Chenyu Zhang, Hongbing Shen

REFERENCES

1. Ludwig JA, Weinstein JN: Biomarkers in cancer staging, prognosis and treatment selection. Nat Rev Cancer 5:845-856, 2005
2. Chen HY, Yu SL, Chen CH, et al: A five-gene signature and clinical outcome in non-small-cell lung cancer. N Engl J Med 356:11-20, 2007
3. Potti A, Mukherjee S, Petersen R, et al: A genomic strategy to refine prognosis in early-stage non-small-cell lung cancer. N Engl J Med 355:570-580, 2006
4. Tsao MS, Sakurada A, Cutz JC, et al: Erlotinib in lung cancer - molecular and clinical predictors of outcome. N Engl J Med 353:133-144, 2005
5. Esquela-Kerscher A, Slack FJ: Oncomirs - microRNAs with a role in cancer. Nat Rev Cancer 6:259-269, 2006
6. Calin GA, Croce CM: MicroRNA signatures in human cancers. Nat Rev Cancer 6:857-866, 2006
7. Ambros V: The functions of animal microRNAs. Nature 431:350-355, 2004
8. Bartel DP: MicroRNAs: Genomics, biogenesis, mechanism, and function. Cell 116:281-297, 2004
9. Lu J, Getz G, Miska EA, Alvarez-Saavedra E, et al: MicroRNA expression profiles classify human cancers. Nature 435:834-838, 2005
10. Yanaihara N, Caplen N, Bowman E, et al: Unique microRNA molecular profiles in lung cancer diagnosis and prognosis. Cancer Cell 9:189-198, 2006
11. Yu SL, Chen HY, Chang GC, et al: MicroRNA signature predicts survival and relapse in lung cancer. Cancer Cell 13:48-57, 2008
12. Chen X, Ba Y, Ma L, et al: Characterization of microRNAs in serum: A novel class of biomarkers for diagnosis of cancer and other diseases. Cell Res 18:997-1006, 2008
13. Resnick KE, Alder H, Hagan JP, et al: The detection of differentially expressed microRNAs from the serum of ovarian cancer patients using a novel real-time PCR platform. Gynecol Oncol 112: 55-59, 2009
14. Gilad S, Meiri E, Yogev Y, et al: Serum microRNAs are promising novel biomarkers. PLoS ONE 3:e3148, 2008
15. Mitchell PS, Parkin RK, Kroh EM, et al: Circulating microRNAs as stable blood-based markers for cancer detection. Proc Natl Acad Sci U S A 105: 10513-10518, 2008
16. Hu Z, Chen J, Tian T, et al: Genetic variants of miRNA sequences and non-small cell lung cancer survival. J Clin Invest 118:2600-2608, 2008
17. Lossos IS, Czerwinski DK, Alizadeh AA, et al R: Prediction of survival in diffuse large-B-cell lymphoma based on the expression of six genes. N Engl J Med 350:1828-1837, 2004
18. Kumar MS, Erkeland SJ, Pester RE, et al: Suppression of non-small cell lung tumor development by the let-7 microRNA family. Proc Natl Acad Sci U S A 105:3903-3908, 2008
19. 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 64:3753-3756, 2004
20. Arriagada R, Bergman B, Dunant A, et al: Cisplatin-based adjuvant chemotherapy in patients with completely resected non-small-cell lung cancer. N Engl J Med 350:351-360, 2004
21. Kato H, Ichinose Y, Ohta M, et al: A randomized trial of adjuvant chemotherapy with uracil-tegafur for adenocarcinoma of the lung. N Engl J Med 350:1713-1721, 2004
22. Winton T, Livingston R, Johnson D, et al: Vinorelbine plus cisplatin vs. observation in resected non-small-cell lung cancer. N Engl J Med 352:2589-2597, 2005
23. Bhattacharjee A, Richards WG, Staunton J, et al: Classification of human lung carcinomas by mRNA expression profiling reveals distinct adenocarcinoma subclasses. Proc Natl Acad Sci U S A 98:13790-13795, 2001
24. Wigle DA, Jurisica I, Radulovich N, et al: Molecular profiling of non-small cell lung cancer and correlation with disease-free survival. Cancer Res 62:3005-3008, 2002
25. Beer DG, Kardia SL, Huang CC, et al: Gene-expression profiles predict survival of patients with lung adenocarcinoma. Nat Med 8:816-824, 2002
26. Endoh H, Tomida S, Yatabe Y, et al: Prognostic model of pulmonary adenocarcinoma by expression profiling of eight genes as determined by quantitative real-time reverse transcriptase polymerase chain reaction. J Clin Oncol 22:811-819, 2004
27. Shedden K, Taylor JM, Enkemann SA, et al: Gene expression-based survival prediction in lung adenocarcinoma: A multi-site, blinded validation study. Nat Med 14:822-827, 2008
28. Lawrie CH, Gal S, Dunlop HM, et al: Detection of elevated levels of tumour-associated microRNAs in serum of patients with diffuse large B-cell lymphoma. Br J Haematol 141:672-675, 2008
29. Skog J, Würdinger T, van Rijn S, et al: Glioblastoma microvesicles transport RNA and proteins that promote tumour growth and provide diagnostic biomarkers. Nat Cell Biol 10:1470-1476, 2008.
30. Valadi H, Ekström K, Bossios A, et al: Exosome-mediated transfer of mRNAs and microRNAs is a novel mechanism of genetic exchange between cells. Nat Cell Biol 9:654-659, 2007
31. Théry C, Zitvogel L, Amigorena S: Exosomes: Composition, biogenesis and function. Nat Rev Immunol 2:569-579, 2002
32. Mack M, Kleinschmidt A, Brühl H, et al: Transfer of the chemokine receptor CCR5 between cells by membrane-derived microparticles: A mechanism for cellular human immunodeficiency virus 1 infection. Nat Med 6:769-775, 2000
33. Al-Nedawi K, Meehan B, Micallef J, et al: Intercellular transfer of the oncogenic receptor EGFRvIII by microvesicles derived from tumour cells. Nat Cell Biol 10:619-624, 2008
34. Nasser MW, Datta J, Nuovo G, et al: Down-regulation of micro-RNA-1 (miR-1) in lung cancer: Suppression of tumorigenic property of lung cancer cells and their sensitization to doxorubicin-induced apoptosis by miR-1. J Biol Chem 283:33394-33405, 2008
35. Datta J, Kutay H, Nasser MW, et al: Methylation mediated silencing of MicroRNA-1 gene and its role in hepatocellular carcinogenesis. Cancer Res 68:5049-5058, 2008
36. Ramalingam SS, Parise RA, Ramanathan RK, et al: Phase I and pharmacokinetic study of vorinostat, a histone deacetylase inhibitor, in combination with carboplatin and paclitaxel for advanced solid malignancies. Clin Cancer Res 13:3605-3610, 2007
37. Sokol NS, Ambros V: Mesodermally expressed Drosophila microRNA-1 is regulated by Twist and is required in muscles during larval growth. Genes Dev 19:2343-2354, 2005
38. Qian S, Ding JY, Xie R, et al: MicroRNA expression profile of bronchioalveolar stem cells from mouse lung. Biochem Biophys Res Commun 377:668-673, 2008
39. Kim CF, Jackson EL, Woolfenden AE, et al: Identification of bronchioalveolar stem cells in normal lung and lung cancer. Cell 121:823-835, 2005
40. Eramo A, Lotti F, Sette G, et al: Identification and expansion of the tumorigenic lung cancer stem cell population. Cell Death Differ 15:504-514, 2008
41. Simon R, Radmacher MD, Dobbin K, et al: Pitfalls in the use of DNA microarray data for diagnostic and prognostic classification. J Natl Cancer Inst 95:14-18, 2003

Acknowledgment

We thank Qingyi Wei, MD, PhD, of The University of Texas M. D. Anderson Cancer Center for critical review and editing of the manuscript.

Appendix

Patients and Methods Study Population

This study was approved by the institutional review board of Nanjing Medical University, and the written informed consent was obtained from each participant or from the patient representatives. Briefly, patients with histopathologically diagnosed incident NSCLC were recruited prospectively onto an ongoing study at the Cancer Hospital of Jiangsu Province (Nanjing) and the First Affiliated Hospital of Nanjing Medical University, Nanjing, China, from July 2003 to July 2008. All patients were genetically unrelated, ethnic Han Chinese, and each participant donated

5 mL of venous blood at their first admission to the hospitals. Exclusion criteria included self-reported previous cancer history, metastasis from other organs, and radiotherapy or chemotherapy before blood collection. The follow-up was conducted every 3 months by telephone calls after the first visit to the hospital.

RNA Isolation and Solexa Sequencing

Briefly, the whole blood was separated into serum and cellular fractions within 24 hours after sample collection. Total RNA was isolated by using the Trizol Reagent (Invitrogen, Carlsbad, CA) according to the manufacturer's instructions. For Solexa sequencing, we used two pooled serum samples; one was pooled from the 30 patients in the longer-survival group and the other was from the 30 patients in the shorter-survival group. Small RNA molecules less than 30 bases were ligated with a pair of Solexa adaptors, amplified by using the adaptor primers for 17 cycle. Sequencing analysis was performed by using the Illumina's Solexa Sequencer (Illumina, San Diego, CA). Finally, clean readouts were compared with the miRBase database (http://microrna.sanger.ac.uk, release 12.0), and the total copy number of the two pooled samples (equal volume) was normalized against an add-on plant miR168.

Quantitative Reverse Transcriptase Polymerase Chain Reaction Assay

All real-time reactions, including no-template controls, were run by using the ABI7300 Real-Time PCR System (Applied Biosystems, Foster City, CA) and were performed in triplicate. Primers used are listed in Appendix Table A1 (online only).

Confirmation of Quantitative RT-PCR Products by Sequencing

We used the forward primers and universal reverse primer to amplify cDNA templates of the four miRNAs (ie, miR-486, miR-30d, miR-1, and miR-499), cloned the polymerase chain reaction (PCR) products into pMD18-T Vector (TaKaRa, Japan) and sequenced them by using BcaBEST sequencing primer M13-47 and RV-M (TaKaRa, Japan).

Statistical Analysis

Differences in demographic and clinical characteristics and mean expression levels of miRNAs were evaluated by χ^2 tests or the *t* test between patient subgroups. Survival time was calculated from the date of lung cancer diagnosis to the date of death or the last follow-up. The associations between overall survival and serum miRNA expression levels were estimated by using the Kaplan-Meier method and log-rank test. Univariate or multivariate Cox proportional hazard regression models were preformed to estimate the crude hazard ratio (HR) or adjusted HRs. Those high-expression miRNAs with HR less than 1 were defined as protective miRNAs, whereas those with HRs greater than 1 were defined as risky miRNAs.

The intensity value was classified and coded as Q1, Q2, Q3, and Q4 according to the 25th, 50th, and 75th percentiles, respectively, of the total miRNA expression levels as cutoff points. In addition, we divided patients in the training set into high-risk and low-risk groups by using the median miRNA expression levels as cutoff points. For the combined analyses, we categorized the four dichotomized miRNAs into a new variable of 0 to 1, 2, 3, and 4 according to the number of high-risk miRNAs (and, for protective miRNAs, the lower expression was categorized as high-risk). The threshold value for each miRNA and risk miRNA categories derived from the training cohort were applied directly to the testing cohort and the combined 303 samples, to see the consistency of the predictive effect of the identified miRNAs.

To additionally investigate the effectiveness of the four-miRNA signature for survival prediction, a mathematical formula was constructed, taking into account both the strength and the positive or negative association of each miRNA with survival. Specifically, we assigned each patient a risk score according to a linear combination of the expression level of the miRNA, weighted by the regression coefficient from the training samples. The risk score was calculated as follows: risk score = (0.969 × expression level of miR-486) + (0.973 × expression level of miR-30d) + (−0.650 × expression level of miR-1) + (−0.815 × expression level of miR-499). Patients were classified as having a high-risk miRNA signature or a low-risk miRNA signature, with the median of the risk score as the threshold value (median, 0.415; range, −4.001 to 3.409). The risk score coefficient and the threshold value derived from the training cohort were also applied directly to the testing cohort and the combined 303 samples. To test the reliability of the threshold value and the related HR of the four-miRNA signature on survival prediction, we conducted 1,000 times of the cross-validation procedure, generated risk-score formulas, generated threshold values on 120 randomly selected participants from the training and testing sets, and calculated HRs and empiric 95% CIs on the entire set of 303 patients.

The Cox stepwise regression model was conducted to determine predictive factors of NSCLC survival, with a significance level of .05 for entering and .10 for removing the respective explanatory variables, and the four miRNAs, age, sex, tumor stage, and histologic type were used as covariates. Stratification analyses of the predictive effect of four-miRNA signature on NSCLC survival were performed by stage and histologic types to avoid the confounding effect from these variables.

All the statistical analyses were performed with Statistical Analysis System software (version 9.1.3; SAS Institute, Cary, NC). A *P* value of less than .05 was considered statistically significant, and all tests were two tailed.

Table A1. miRNA Primer Information

Name	Forward Primer (5′-3′)	Reverse Primer (5′-3′)
hsa-let-7a	ACACTCCAGCTGGGTGAGGTAGTAGGTTGT	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAACTATAC
hsa-let-7g	ACACTCCAGCTGGGTGAGGTAGTAGTTTGT	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAACTGTAC
hsa-let-7i	ACACTCCAGCTGGGTGAGGTAGTAGTTTGT	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAACAGCAC
hsa-miR-1	ACACTCCAGCTGGGTGGAATGTAAAGAAGT	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGATACATAC
hsa-miR-21	ACACTCCAGCTGGGTAGCTTATCAGACTGA	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGTCAACATC
hsa-miR-22	ACACTCCAGCTGGGAAGCTGCCAGTTGAAG	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGACAGTTCT
hsa-miR-26b	ACACTCCAGCTGGGTTCAAGTAATTCAGG	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGACCTATCC
hsa-miR-30d	ACACTCCAGCTGGGTGTAAACATCCCCGAC	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGCTTCCAGT
hsa-miR-146b	ACACTCCAGCTGGGTGAGAACTGAATTCCA	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAGCCTATG
hsa-miR-206	ACACTCCAGCTGGGTGGAATGTAAGGAAGT	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGCCACACAC
hsa-miR-378	ACACTCCAGCTGGGACTGGACTTGGAGTC	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGCCTTCTGA
hsa-miR-486	ACACTCCAGCTGGGTCCTGTACTGAGCTGC	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGCTCGGGGC
hsa-miR-499	ACACTCCAGCTGGGTTAAGACTTGCAGTG	CTCAACTGGTGTCGTGGAGTCGGCAATTCAGTTGAGAAACATCA
Universal reverse primer		TGGTGTCGTGGAGTCG

Abbreviation: miRNA, microRNA.

Table A2. Clinical and Demographic Information of the Data Sets

	Discovery Set (n = 60)									
	Long Survival Group (n = 30)		Short Survival Group (n = 30)			Training Set (n = 120)		Testing Set (n = 123)		
Characteristic	No.	%	No.	%	*P**	No.	%	No.	%	*P**
Sex					.542					.014
Male	22	73.3	24	80.0		78	65.0	98	79.7	
Female	8	26.7	6	20.0		42	35.0	25	20.3	
Age, years					.658					.139
Mean	60.3		59.4			58.6		60.4		
SD	6.94		8.63			9.44		9.46		
Smoking status					1.000					.296
Nonsmoker	9	30.0	9	30.0		51	42.5	44	35.8	
Ever-smoker	21	70.0	21	70.0		69	57.5	79	64.2	
Histologic type					.796					.186
Squamous cell carcinoma	14	46.7	15	50.0		49	40.8	40	32.5	
Adenocarcinoma	16	53.3	15	50.0		71	59.2	83	67.5	
Stage					1.000					.471
I	10	33.3	10	33.3		36	30.0	46	37.4	
II	10	33.3	10	33.3		21	17.5	20	16.3	
IIIa	10	33.3	10	33.3		63	52.5	57	46.3	

Abbreviation: SD, standard deviation.
*Two-sided χ^2 test or Student *t* test.

Table A3. Expression Levels of the Four miRNAs and Survival of Patients With NSCLC in the Discovery Data Set

miRNA	Patients (N = 60)	Deaths (n = 30)	MST (months)	Log-Rank *P*	Crude Analysis HR	Crude Analysis 95% CI	Adjusted Analysis HR	Adjusted Analysis 95% CI	Trend *P*
miR-486									
Q1, ≤ 0.470	14	3	Not reached		1.00		1.00		
Q2, 0.470-0.830	16	5	Not reached		1.68	0.40 to 7.03	1.81	0.43 to 7.68	
Q3, 0.830-1.190	15	10	9.97		4.94	1.35 to 18.05	5.28	1.42 to 19.66	
Q4, > 1.190	15	12	13.77	.009	5.07	1.42 to 18.12	5.47	1.48 to 20.19	.001
miR-30d									
Q1, ≤ 0.325	15	1	Not reached		1.00		1.00		
Q2, 0.325-0.760	15	5	Not reached		6.10	0.71 to 52.23	6.35	0.73 to 55.18	
Q3, 0.760-1.280	15	10	13.77		14.36	1.83 to 112.75	16.70	2.08 to 133.93	
Q4, > 1.280	15	14	8.7	< .001	29.28	3.81 to 225.35	38.59	4.89 to 304.78	< .001
miR-1									
Q4, > 1.465	15	1	Not reached		1.00		1.00		
Q3, 0.525-1.465	15	5	Not reached		6.16	0.72 to 52.73	6.01	0.69 to 52.20	
Q2, 0.180-0.525	15	9	12.27		13.84	1.75 to 109.52	13.90	1.72 to 112.49	
Q1, ≤ 0.180	15	15	9.8	< .001	50.80	6.47 to 398.86	54.77	6.80 to 441.47	.003
miR-499									
Q4, > 1.115	15	2	Not reached		1.00		1.00		
Q3, 0.815-1.115	15	6	Not reached		3.46	0.70 to 17.13	3.32	0.66 to 16.60	
Q2, 0.605-0.815	15	11	15.17		8.89	1.96 to 40.41	9.45	2.04 to 43.80	
Q1, ≤ 0.605	15	11	7.2	.001	9.54	2.11 to 43.18	8.76	1.89 to 40.71	< .001

NOTE. Analysis was adjusted for age, sex, and histology.
Abbreviations: miRNA, microRNA; NSCLC, non–small-cell lung cancer; MST, median survival time; HR, hazard ratio; Q, quartile.

Table A4. Four miRNA Expressions and Survival of Patients With NSCLC in the Training and Testing Sets

miRNA	No. of Patients Training	No. of Patients Testing	No. of Deaths Training	No. of Deaths Testing	MST (months) Training	MST (months) Testing	Log-Rank *P* Training	Log-Rank *P* Testing	Training Adjusted Analysis HR	Training Adjusted Analysis 95% CI	Testing Adjusted Analysis HR	Testing Adjusted Analysis 95% CI
miR-486												
Low, ≤ 0.795	60	62	16	20	Not reached	39.93			1.00		1.00	
High, > 0.795	60	61	24	23	31.33	31.33	.070	.058	1.86	0.94 to 3.66	1.71	0.92 to 3.20
miR-30d												
Low, ≤ 0.680	60	58	13	17	Not reached	42.6			1.00		1.00	
High, > 0.680	60	65	27	26	22.97	31.6	< .001	.007	4.46	2.17 to 9.17	2.97	1.51 to 5.86
miR-1												
High, > 0.675	60	62	11	16	41.67	42.6			1.00		1.00	
Low, ≤ 0.675	60	61	29	27	22.97	31.8	.006	.021	2.95	1.44 to 6.06	3.11	1.57 to 6.19
miR-499												
High, > 0.750	58	74	13	21	38.53	43.17			1.00		1.00	
Low, ≤ 0.750	62	49	27	22	22.97	31.6	.005	.049	2.61	1.33 to 5.15	2.44	1.27 to 4.70
No. of high-risk miRNAs												
0-1	37	41	4	10	Not reached	43.17			1.00		1.00	
2	43	48	8	12	41.67	36.77			2.32	0.68 to 7.96	2.06	0.81 to 5.20
3	34	30	23	17	17.83	18.73			17.42	5.58 to 54.36	13.06	4.84 to 35.22
4	6	4	5	4	17.03	17.08	< .001	< .001	19.57	4.94 to 77.50	19.20	5.15 to 71.55

NOTE. Analysis was adjusted for age, sex, smoking status, histology, and stage.
Abbreviations: miRNA, microRNA; NSCLC, non–small-cell lung cancer; MST, median survival time; HR, hazard ratio.

Table A5. Stepwise Cox Proportional Hazards Analyses on NSCLC Survival

Final Variable	β	SE	HR	95% CI	*P*
Stage	.51	0.12	1.67	1.32 to 2.12	< .001
miR-486, high *v* low	.90	0.22	2.47	1.61 to 3.79	< .001
miR-30d, high *v* low	1.41	0.23	4.10	2.62 to 6.42	< .001
miR-1, low *v* high	1.17	0.23	3.22	2.04 to 5.08	< .001
miR-499, low *v* high	1.02	0.22	2.78	1.81 to 4.27	< .001
Stage	.48	0.12	1.62	1.29 to 2.03	< .001
The four-miRNA signature, > 0.415 *v* < 0.415	2.23	0.24	9.31	5.79 to 14.97	< .001

NOTE. Analysis was adjusted for age, sex, and smoking status.
Abbreviations: NSCLC, non–small-cell lung cancer; SE, standard error; HR, hazard ratio; miRNA, microRNA.

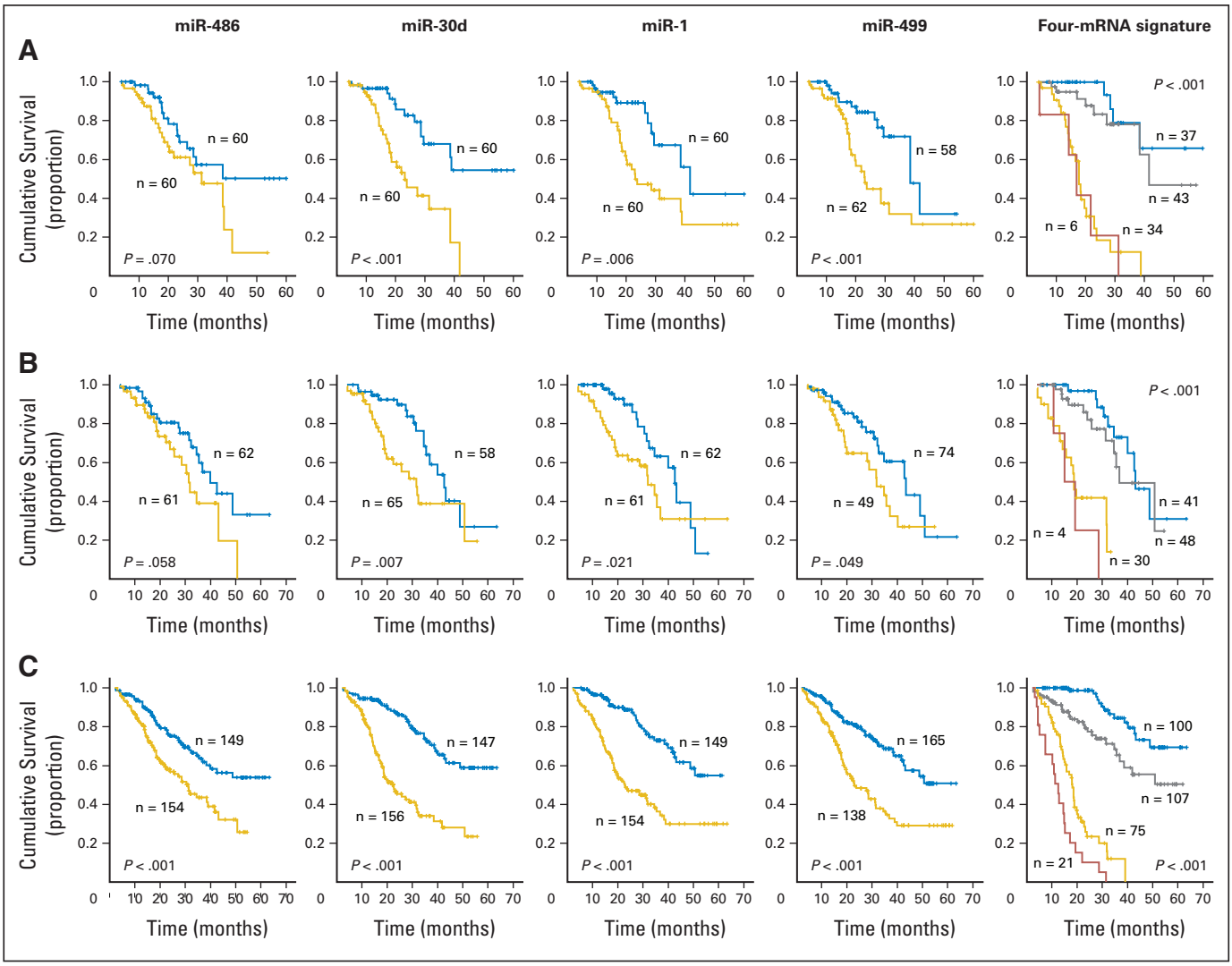

Fig A1. Kaplan-Meier estimates of overall survival of patients with non–small-cell lung cancer according to the four microRNAs (miRNAs) and their combinations. (A) Training set (120 patients); (B) testing set (123 patients); (C) all-combined data sets (303 patients). Labels at the top of the figure apply to all graphs in the same column.

