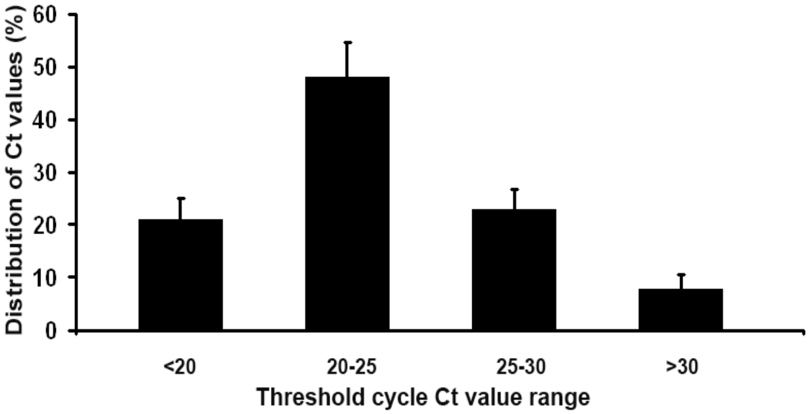
SUPPLEMENT MATERIALS

Supplement Fig. 1.
Expression of the 377 miRNAs were readily detectable in surgical tissue specimens.



Three hundreds and seventy-seven human miRNAs could easily be measured on the surgically resected lung cancerous tissues and noncancerous lung tissues by real-time PCR-based miRNA array. A TaqMan-based miRNA array (Applied Biosystems, Foster City, CA) was used to profile the mature miRNAs in lung adenocarcinoma and normal lung tissues. To determine if expression of these 377 miRNAs was detectable, two RNA pools of malignant and nonmalignant were prepared containing equal amounts of RNA from 20 samples per group and tested by the miRNA array analysis. Of the 377 genes, 346 (92%) had ≤ 30 Ct value, however only 32 (8%) displayed >30 Ct value.

Supplement Table 1. Demographic and clinical characteristics in a case-control cohort of cancer patients and healthy individuals

	Lung adenocarcinoma patients (n = 36)	Controls (n = 36)	P-value
Age	68.2 (SD 8.4)	66.7 (SD 4.9)	0.89
Sex			
Female	16	16	
Male	20	20	
Race			
White	26	26	
African American	10	10	
Smoking status			
Pack-years	28.6 (SD 25.9)	28.8 (SD 28.3)	0.92
Stage	All are stage I		
Histology	All are adenocarcinomas		

Abbreviations: SD, standard deviation.

Supplement Table 2. Annotated miRNAs differentially expressed in lung adenocarcinomas verus normal lung tissues (average fold-change ≥1.0, p<0.01)

MiRNAs	Log2 (fold change)
hsa-miR-486	**-2.968319667**
hsa-miR-145	**-2.660679833**
hsa-miR-126	**-2.156865167**
hsa-miR-138	**-1.784505667**
hsa-miR-139	**-1.735165667**
hsa-miR-143	**-1.700703535**
hsa-miR-218	**-1.661419667**
hsa-miR-150	**-1.618181664**
hsa-miR-140	**-1.511957436**
hsa-miR-195	-1.393894167
hsa-miR-16	-1.299653167
hsa-let-7b	-1.141909167
hsa-miR-222	-1.102005167
hsa-miR-27a	-1.050386578
hsa-miR-886	-1.048786558
hsa-miR-10a	-1.046658586
hsa-miR-376c	-1.035308517
hsa-miR-30c	-1.029413167
hsa-miR-99a	-1.013456167
hsa-miR-30b	-1.000671167
hsa-miR-501	1.066856836
hsa-miR-429	1.186442682
hsa-miR-135b	1.240513588
hsa-miR-196b	1.352369286
hsa-miR-720	**1.508929647**
hsa-miR-92b	**1.538657833**
hsa-miR-130b	**1.567958333**
hsa-miR-31	**1.601236526**
hsa-miR-93	**1.682514576**
hsa-miR-210	**1.747803333**
hsa-miR-200a	**1.862715328**
hsa-miR-182	**2.217326833**
hsa-miR-200b	**2.268104167**
hsa-miR-21	**2.560514667**
hsa-miR-375	**2.891464333**

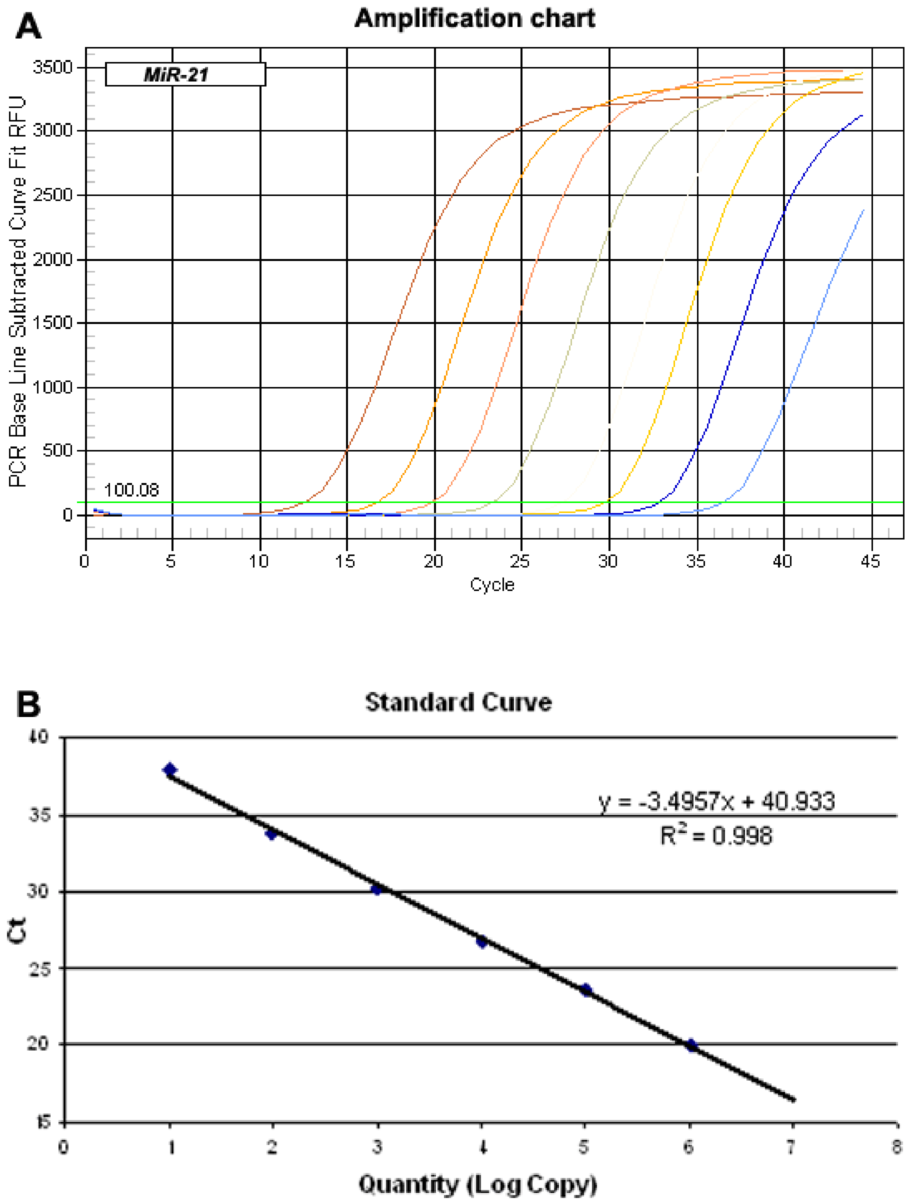
The miRNAs with average fold-change ≥1.5 are given in bold type (Please see the following supplement table 3)

Supplement Table 3. Annotated miRNAs differentially expressed in lung adenocarcinomas verus normal lung tissues (average fold-change ≥1.5, p<0.01)

MiRNAs	Log2 (fold change)
hsa-miR-486	**-2.968319667**
hsa-miR-145	**-2.660679833**
hsa-miR-126	**-2.156865167**
hsa-miR-138	-1.784505667
hsa-miR-139	-1.735165667
hsa-miR-143	-1.700703535
hsa-miR-218	-1.661419667
hsa-miR-150	-1.618181664
hsa-miR-140	-1.511957436
hsa-miR-720	1.508929647
hsa-miR-92b	1.538657833
hsa-miR-130b	1.567958333
hsa-miR-31	1.601236526
hsa-miR-93	1.682514576
hsa-miR-210	1.747803333
hsa-miR-200a	1.862715328
hsa-miR-182	**2.217326833**
hsa-miR-200b	**2.268104167**
hsa-miR-21	**2.560514667**
hsa-miR-375	**2.891464333**

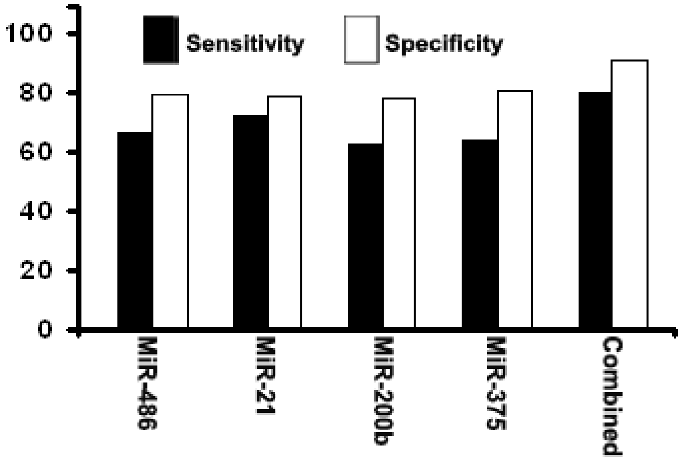
The seven miRNAs with average fold-change ≥2.0 are given in bold type.

Supplement Fig. 2. **Expression of the miRNAs can sensitively be detected in sputum.**



To determine whether the seven newly identified miRNAs can be specifically and reliably detected in sputum specimens. Two RNA pools containing equal amounts of RNA from sputum samples of 36 cancer patients and 36 cancer-free individuals were prepared, respectively. A. To determine the sensitivity of detecting the miRNAs by RT-qPCR in sputum, the total RNA was serially diluted in DEPC water. Amplification plot of total RNA from sputum specimen at seven orders of magnitude was made for the detection of *miR-21*. B. Correlation of total RNA input with the threshold of cycle (Ct) values for *miR-21* assay ($R^2 = 0.998$, slope =–3.496). The assays were performed on the sputum samples for all seven miRNAs. The supplement Fig. 2 only shows the results of *miR-21*. All experiments were done in triplicates.

Supplement Fig. 3. The combination of the four miRNAs display higher specificity and specificity in the identification of lung adenocarcinoma compared with a single one used a lone.



Comparison of four miRNAs individually and the four miRNAs combined in identification of lung adenocarcinoma patients in a case-control cohort of 36 patients diagnosed with stage I lung adenocarcinoma and 36 health subjects. The four miRNAs in combination produced a sensitivity of 80.6% and a sensitivity of 91.7%, being statistically higher than 59.5-72.6% sensitivity and 73.8-82.9% specificity of the individual miRNAs (All $p<0.05$).

Supplement Table 4. Expression levles of the four miRNAs and their diagnostic significance in sputum of 64 NSCLC patients and 58 healthy controls

MiRNA genes	Mean (SD) in NSCLC patients	Mean (SD) in controls	P*	AUC* (SE)	Sensitivity (%)	Specificity (%)
MiR-486	0.049 (0.046)	0.193 (0.168)	<0.001	0.727 (0.030)	62.6	69.4
MiR-21	6.198 (4.891)	2.823 (1.289)	<0.001	0.752 (0.038)	58.8	73.8
MiR-200b	5.368 (4.562)	2.642 (1.873)	<0.001	0.789 (0.027)	55.1	72.2
MiR-375	6.381 (5.001)	2.992 (1.654)	<0.001	0.713 (0.026)	60.3	71.7
All four miRNAs				0.839 (0.049)	70.3	80

Abbreviations: NSCLC: non–small-cell lung cancer; SD, standard deviation; AUC, the area under receiver operating characteristic curve; SE, standard error.

*Both the P value and AUC were obtained using the *U6*-normalized values

Supplement Table 5. Association of the miRNA expressions with clinical and demographic characteristics of the NSCLC patients

	MiRNA, OR (95% CI), P*			
Characteristics	*MiR-486*	*MiR-21*	*MiR-200b*	*MiR-375*
Age	0.66 (0.30 to 1.46), 0.29	0.78 (0.36 to 2.00), 0.59	0.65 (0.27 to 1.48). 0.28	0.67 (0.25 to 1.28), 0.19
Gender	1.65 (0.88 to 3.26), 0.12	1.62 (0.83 to 3.26), 0.18	1.57 (0.82 to 3.13), 0.19	1.34 (0.65 to 2.80), 0.39
Ethnic group	1.39 (0.69 to 2.84), 0.37	1.68 (0.83 to 3.63), 0.18	1.71 (0.34 to 3.87), 0.19	1.35 (0.67 to 2.83), 0.56
Smoking status	1.36 (0.55 to 3.28), 0.57	1.26 (0.48 to 3.45), 0.58	1.40 (0.56 to 3.48), 0.48	1.46 (0.53 to 3.28), 0.49
Tumor histological type	2.52 (1.23 to 5.29), **0.02**	2.70 (1.83 to 5.57), **0.01**	2.55 (0.97 to 4.95), **0.03**	2.37 (1.86 to 4.97), **0.03**
Tumor location	2.91 (1.55 to 6.12), **0.03**	2.64 (1.35 to 5.69), **0.01**	2.63 (0.10 to 5.15), **0.01**	2.20 (1.23 to 4.85), **0.05**
Stage	1.95 (0.99 to 3.78), 0.07	1.67 (0.84 to 3.28), 0.16	1.46 (0.76 to 2.77), 0.28	1.58 (0.84 to 2.96), 0.18

Abbreviations: NSCLC, non-small cell lung cancer; OR = odds ratio; CI = confidence interval.

*Uunivariate analysis. All P values are two-sided, and $P \leq 0.05$ was considered statistically significant and given in bold type.

Supplement Table 6. Association of the miRNA expressions with demographic characteristics of the healthy subjects

	MiRNA, OR (95% CI), P*			
Demographic characteristics	*MiR-486*	*MiR-21*	*MiR-200b*	*MiR-375*
Age	0.62 (0.27 to 1.52), 0.26	0.67 (0.35 to 2.28), 0.62	0.67 (0.26 to 1.62), 0.29	0.69 (0.26 to 1.22), 0.15
Gender	1.68 (0.84 to 3.27), 0.19	1.64 (0.86 to 3.21), 0.16	1.50 (0.86 to 3.37), 0.22	1.36 (0.64 to 2.34), 0.34
Ethnic group	1.32 (0.62 to 2.85), 0.36	1.63 (0.81 to 3.60), 0.17	1.65 (0.32 to 3.85), 0.16	1.38 (0.63 to 2.92), 0.59
Smoking status	1.66 (0.58 to 3.25), 0.64	1.28 (0.39 to 3.49), 0.64	1.48 (0.48 to 3.42), 0.46	1.50 (0.56 to 3.33), 0.45

Abbreviations: OR = odds ratio; CI = confidence interval.

*Uunivariate analysis. All P values are two-sided, and $P \leq 0.05$ was considered statistically significant.