Application of serum microRNA-9-5p, 21-5p, and 223-3p combined with tumor markers in the diagnosis of non-small-cell lung cancer in Yunnan in southwestern China

Yanlong Yang1
Kai Chen1,2
Yongchun Zhou3–5
Zaoxiu Hu6
Shuai Chen1
Yunchao Huang1,3–5

1Department of Thoracic Surgery I, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital, Yunnan Cancer Center), Kunming, 2Department of Cardiothoracic Surgery, The Second Affiliated Hospital of Wenzhou Medical University, Wenzhou, 3Cancer Research Institute of Yunnan Province, 4Key Laboratory of Lung Cancer Research of Yunnan Province, 5International Joint Laboratory on High Altitude Regional Cancer of Yunnan Province, 6Department of Pathology, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital, Yunnan Cancer Center), Kunming, People’s Republic of China

Correspondence: Yunchao Huang
Department of Thoracic Surgery I, The Third Affiliated Hospital of Kunming Medical University (Yunnan Cancer Hospital, Yunnan Cancer Center), Number 519 Kunzhao Road, Xishan District, Kunming 650118, Yunnan Province, People’s Republic of China
Tel +86 871 818 5656
Email hycyn2008@163.com

Purpose: Xuanwei City is located in late Permian coal-accumulating areas of the northeastern region of Yunnan Province. In China, morbidity and mortality from lung cancer are highest in Yunnan. Identifying useful circulating markers suitable for the diagnosis of lung cancer in this region is quite meaningful. In this study, we evaluated diagnostic roles of serum miR-9-5p, 21-5p, 223-3p, 135b-5p, 339-5p, and 501-5p in patients with non-small-cell lung cancer (NSCLC) in Yunnan. Moreover, we evaluated the diagnostic performance of several tumor markers, including carcinoembryonic antigen (CEA), cytokeratin 19 fragment 21-1 (CYFRA21-1), and squamous cell carcinoma-related antigen (SCC).

Methods: Quantitative real-time polymerase chain reaction detected six miRNAs in the serum of 104 NSCLC patients and 50 cancer-free controls. Other markers, including CEA, CYFRA21-1, and SCC, in serum were also measured. The diagnostic ability of miRNAs and tumor markers was evaluated by receiver operating characteristic (ROC) curve analysis. The diagnostic performance of these serum markers was also evaluated in Xuanwei and non-Xuanwei subjects, because the etiological and the epidemiological characteristics of lung cancer in Xuanwei were quite different from those in other regions.

Results: Serum miR-9-5p, miR-21-5p, miR-223-3p, CEA, CYFRA21-1, and SCC were upregulated in NSCLC patients, compared with cancer-free controls. No significant difference was found in miR-135b-5p, miR-339-5p, and miR-501-5p expression. The area under ROC curves (AUCs) of miR-9-5p, miR-21-5p, miR-223-3p, CEA, CYFRA21-1, and SCC were 0.706, 0.765, 0.744, 0.749, 0.735, and 0.616, respectively. When combined, miRNAs and tumor markers yielded the highest diagnostic power, with AUC of 0.886, sensitivity of 82.69%, and specificity of 88.00%. In Xuanwei subjects, miR-223-3p and CEA may be suitable biomarkers to distinguish NSCLC from cancer-free states with AUCs of 0.752 and 0.791, respectively. The diagnostic power of the combination of miRNAs and tumor markers was still the highest in both subgroups (region: Xuanwei and non-Xuanwei; stages: I–II and III–IV).

Conclusion: Serum miR-9-5p, miR-21-5p, miR-223-3p, CEA, CYFRA21-1, and SCC could be potential diagnostic biomarkers for NSCLC patients in Yunnan. miRNAs and tumor markers should be combined to diagnose NSCLC, as it showed better ability for screening patients with NSCLC.

Keywords: non-small-cell lung cancer, microRNA, tumor markers, diagnosis, biomarker

Introduction

Currently, the mortality rate of lung cancer is still the highest among all cancers in both men and women worldwide.1 Data indicated that an estimated 733,300 incident lung
cancer cases and 610,200 lung cancer deaths would occur in China in 2015.2 Xuanwei City is located in late Permian coal-accumulating areas in the northeastern regions of the Yunnan province. The incidence and mortality rate of lung cancer are highest in China. In 1973–1975, 1990–1992, and 2004–2005, standardized lung cancer death rates in Xuanwei were 28.2, 40.29, and 83.28 per 100,000 respectively; however, in the China, they were 5.6, 15.19, and 20.42 per 100,000, respectively. The standardized lung cancer death rate in Yunnan was twice that in China. Two major noticeable features of lung cancer in Xuanwei were: 1) the incidence and mortality rate of lung cancer in females were rather high and almost all of them were nonsmokers; and 2) the major type of lung cancer in women was adenocarcinoma. Previous studies have indicated that Xuanwei is located in the late Permian coal-accumulating areas that are rich in bituminous (smoky) coal. The main reason for the high incidence and mortality of lung cancer is attributed to indoor air pollution caused by the use of “smoky coal” in unvented indoor fire pits. Burning of “smoky coal” releases carcinogenic substances such as polycyclic aromatic hydrocarbons, particulate matter, and crystalline quartz. A stove improvement project was undertaken in the late 1980s; the government encouraged and supported local residents to use stoves with chimneys. Even after changing stoves, the mortality rate due to lung cancer did not decline as expected. As a result, the earlier detection of lung cancer in Yunnan, especially in Xuanwei, would be greatly meaningful to improve the prognosis of this lethal disease.

One of the major challenges in lung cancer research is the identification of stable biomarkers that can be routinely measured. Several serum biomarkers such as carcinoembryonic antigen (CEA), cytokeratin 19 fragment 21-1(CYFRA21-1), and squamous cell carcinoma-related antigen (SCC) for lung cancer have been widely used in recent decades. In recent years, more attention has been paid to circulating microRNAs (miRNAs). miRNAs are short, conserved, non-coding RNAs that could negatively regulate gene expression at the post-transcriptional level by binding the 3’-untranslated region of target mRNAs, resulting in either mRNA degradation or translational repression. Passively leaked or actively transported from cells, circulating miRNAs could be stably detected in blood and have been used as biomarkers for diagnosis, prognosis, or monitoring the curative effect in various cancers, including lung cancer.

Here, we intended to investigate whether serum miR-9-5p, 21-5p, 223-3p, 135b-5p, 339-5p, and 501-5p were suitable for use as diagnostic biomarkers for patients with non-small-cell lung cancer (NSCLC) in Yunnan. miR-9-5p, 223-3p, 135b-5p, and 501-5p were identified from a microarray result that compared 24 NSCLC patients and their paired adjacent normal tissues in our previous work (unpublished data). All of these four miRNAs were upregulated (for miR-9-5p, fold change [FC] =4.877, \( P=0.026 \); for miR-223-3p, FC =3.61, \( P=0.00008 \); for miR-135b-5p, FC =8.00, \( P=0.0007 \); and, for miR-501-5p, FC =3.17, \( P=0.022 \)), and were reported to act as oncomirs in the development of cancer. Both miR-21-5p and miR-339-5p were identified from the literature. miR-21-5p has been widely investigated in various cancers. miR-339-5p is downregulated in colorectal cancer, hepatocellular carcinoma, and breast cancer. We tested the expression level of serum miR-9-5p, 21-5p, 223-3p, 135b-5p, 339-5p, and 501-5p in NSCLC and cancer-free subjects. Differentially expressed miRNAs were further analyzed for their diagnostic roles. In addition, we combined miRNAs with CEA, CYFRA21-1, and SCC to generate better diagnostic performance in the diagnosis of NSCLC in Yunnan.

Materials and methods

Patients and control individuals

Serum samples of 104 NSCLC patients were collected between October 2016 and May 2017 in the Department of Thoracic Surgery I, The Third Affiliate Hospital of Kunming Medical University. All 104 patients had pathologically confirmed NSCLC and had not undergone any therapeutic procedure before admission to the hospital. The assessment of disease stage was based on the seventh edition of the tumor–node–metastasis (TNM) staging system of the International Association for the Study of Lung Cancer. Additionally, in the same period, serum samples were collected from 50 cancer-free individuals without a history of malignancy. Written informed consent was obtained from all included individuals, and this study was approved by the ethical committee of The Third Affiliated Hospital of Kunming Medical University. All 154 subjects were from the Yunnan province. There was no difference in age, sex, smoking status, and the region (Xuanwei or non-Xuanwei) between NSCLC patients and the cancer-free group. Detailed characteristic of the two study groups are presented in Table 1. Blood samples of NSCLC patients and cancer-free controls were collected into 10 mL tubes without anticoagulant. Then, samples were stored at room temperature for a few minutes and centrifuged at 3,000 rpm for 10 minutes. Supernatant serum samples were then carefully transferred...
to new RNA-free Eppendorf tubes and stored at $-80^\circ$C until further use.

**RNA isolation, reverse transcription, and quantitative real time polymerase chain reaction (qRT-PCR)**

The miRNeasy serum/plasmakit (QIAGEN, Hilden, Germany) was used to extract miRNA according to the manufacturer’s protocol. Total RNA (100 ng) was reverse transcribed into cDNA using miRcute Plus miRNA First-Strand cDNA Synthesis Kit (Tiangen Biotech, Beijing, China) in a T100™ Thermal Cycler (Bio-Rad, Pleasanton, CA, USA). The miRcute Plus miRNA qPCR Detection Kit (Tiangen Biotech) was applied for qRT-PCR in a ABI 7900HT Fast Real Time PCR System (Thermo Fisher Scientific, Waltham, MA, USA), according to manufacturer’s instructions. Primers of miR-9-5p (#CD201-0142), miR-21-5p (#CD201-0092), miR-223-3p (#CD201-0099), miR-135b-5p (#CD201-0213), miR-339-5p (#CD201-0360), and miR-501-5p (#CD201-0641) were obtained from Tiangen Biotech. All reaction mixtures were incubated in a 96-well optical plate at 95°C for 15 minutes, followed by 40 cycles at 94°C for 20 seconds, and at 60°C for 34 seconds. Cycle threshold (Ct) values correspond to the number of cycles required for the amplified product to reach a critical threshold of detection. The expression level of studied miRNAs was normalized relative to miR-16, and that of each miRNA was quantified through normalized threshold Ct, wherein

$$\Delta Ct = [\text{Ct (miRNA)}] - [\text{Ct (U6)}]$$

and the relative expression level was calculated as $2^{-\Delta Ct}$.23

**Serum tumor marker assessment**

Fasting peripheral vein blood specimens were collected from subjects. For each group, specimens were collected on the day after subjects were admitted to the hospital. None of them underwent any therapeutic procedures prior to hospitalization. Levels of CEA, CYFRA21-1, and SCC were measured in serum specimens at the Clinical Laboratory of The Third Affiliate Hospital of Kunming Medical University. CEA and CYFRA21-1 levels were measured by electrochemiluminescent immunoassay (ECLIA; Hoffman-La Roche Ltd., Basel, Switzerland). The SCC level was measured by chemiluminescent microparticle immunoassay (CMIA, ARCHITECT SCC assay; Abbott Laboratories, Abbott Park, IL, USA).

**Statistical analysis**

Categorical data were expressed as frequencies. Pearson’s chi-square test or Fisher’s exact test was used to analyze differences in categorical variables. Student’s $t$-test or the Mann–Whitney $U$-test was used to analyze differences in continuous variables between the two groups. SPSS 19.0 (SPSS Inc., Chicago, IL, USA) was applied for data analysis. Graphical displays were prepared by using GraphPad Prism 5 (GraphPad Software, Inc, La Jolla, CA, USA) to show distributions of the expression for each serum miRNAs and tumor markers in each study group. MedCalc version 12 (MedCalc Software, Mariakerke, Belgium) was used to assess diagnostic accuracy; receiver operating curves (ROCs) with area under curves (AUCs) were generated.24 Values for specificity, sensitivity, positive/negative likelihood ratio (+LR/−LR), and Youden’s index (YI) for each factor and the combined groups were used to determine the cutoff value for each biomarker’s concentration, and with these we could distinguish between patients and controls. In each analysis, $P$-values of $<0.05$ were assessed as being statistically significant.

| Table 1 Baseline characteristics of study subjects |
|-----------------|-----------------|-----------------|
| **Variable**    | **Cases** (n=104) | **Controls** (n=50) | **P-value** |
| Age (years)     | 57.03±9.52      | 55.06±8.47      | 0.215       |
| Sex             |                 |                 |             |
| Male            | 59              | 22              | 0.138       |
| Female          | 45              | 28              |             |
| Smoking         |                 |                 |             |
| Yes             | 58              | 21              | 0.109       |
| No              | 46              | 29              |             |
| Region          |                 |                 |             |
| Xuanwei         | 37              | 16              | 0.149       |
| Non-Xuanwei     | 67              | 34              |             |
| Histology       |                 |                 |             |
| LUAD            | 59              | –               |             |
| LUSC            | 40              | –               |             |
| Others          | 5               | –               |             |
| Stage           |                 |                 |             |
| I               | 36              | –               |             |
| II              | 11              | –               |             |
| III             | 35              | –               |             |
| IV              | 22              | –               |             |
| Subtypes of benign diseases |             |                 |             |
| Pulmonary infection | –           | 16              |             |
| Inflammatory pseudotumor | –     | 12              |             |
| Hamartoma       | –               | 8               |             |
| Pulmonary tuberculosis | –          | 12              |             |
| Others          | –               | 2               |             |

**Abbreviations:** –, not applicable; LUAD, lung adenocarcinoma; LUSC, lung squamous cell carcinoma.
Results
Altered miR-9-5p, miR-21-5p, miR-501-5p, miR-223-5p, miR-135b-5p, and miR-339-5p expression and their clinicopathological association in NSCLC

All assays of samples from 104 NSCLC patients and 50 cancer-free individuals were successfully conducted on serum specimens. As shown in Figure 1, expressions of serum miR-9-5p, miR-21-5p, and miR-223-3p were significantly higher in NSCLC patients than in cancer-free patients. However, serum miR-135b-5p, 339-5p, and 501-5p expressions showed no significant difference between the two groups. As shown in Table 2, high miR-9-5p expression was associated with positive lymphatic metastasis ($P=0.045$) and distal metastasis ($P=0.045$). The expression level of miR-21-5p in lung squamous cell carcinoma (LUSC) was higher than in lung adenocarcinoma (LUAD) patients ($P=0.009$). Furthermore, high miR-21 expression was associated with advanced stage ($P=0.02$) and T factor ($P=0.028$). No significant association was found between miR-223-3p expression and clinicopathological parameters in NSCLC patients.

Altered CEA, CYFRA21-1, and SCC levels and their clinicopathological association in NSCLC

Levels of CEA, CYFRA21-1, and SCC were higher in NSCLC patients than in cancer-free subjects (Figure 1). In LUSC cases, CYFRA21-1 and SCC levels were higher than in LUAD patients. Patients with T2–4 tumors had higher CEA, CYFRA21-1, and SCC levels than T1 patients. Higher CEA, CYFRA21-1, and SCC levels were associated with positive lymphatic metastasis and distal metastasis. Patients with advanced stages (III–IV) had higher CYFRA21-1 and SCC levels than those with early-stage (I–II) cancer (Table 2).

Diagnostic value of miR-9-5p, miR-21-5p, miR-223-3p, CEA, CYFRA21-1, and SCC in overall patients

An ROC analysis was conducted in three differentially expressed miRNAs (miR-9, miR-21, and miR-223) and three tumor markers (CEA, CYFRA21-1, and SCC). The ROC curves are shown in Figure 2 and results are summarized in Tables 3 and 4. The ROC analysis showed that miR-9-5p
| Variable                        | N   | miR-9-5p* | miR-21-5p* | miR-223-3p* | CEA*   | CYFRA21-1* | SCC*   |
|--------------------------------|-----|-----------|------------|-------------|--------|------------|--------|
| Sex                            |     |           |            |             |        |            |        |
| Male                           | 59  | 0.080     | 3.160      | 3.758       | 3.980  | 3.500      | 1.100  |
| Female                         | 45  | 0.090     | 3.784      | 4           | 5.000  | 3.000      | 0.500  |
| P-values*                      | 0.570| 0.937     | 0.911      | 0.260       | 0.126  | <0.001     |        |
| Smoking status                 |     |           |            |             |        |            |        |
| Yes                            | 58  | 0.082     | 3.160      | 3.556       | 4.390  | 3.6        | 1.000  |
| No                             | 46  | 0.085     | 3.657      | 4.17        | 4.440  | 3 (2.5, 5.375) | 0.700  |
| P-values*                      | 0.756| 0.997     | 0.344      | 0.906       | 0.248  | 0.001      |        |
| Region                         |     |           |            |             |        |            |        |
| Xuanwei                        | 37  | 0.056     | 3.160      | 3.758       | 4.350  | 2.80       | 0.700  |
| Non-Xuanwei                    | 67  | 0.095     | 3.340      | 3.864       | 4.440  | 3.500      | 1.000  |
| P-values*                      | 0.056| 0.699     | 0.573      | 0.615       | 0.060  | 0.089      |        |
| Histology                      |     |           |            |             |        |            |        |
| LUAD                           | 59  | 0.085     | 2.707      | 3.771       | 4.550  | 3.000      | 0.700  |
| LUSC                           | 40  | 0.082     | 5.044      | 4.363       | 4.180  | 4.600      | 1.700  |
| P-values*                      | 0.921| 0.009     | 0.632      | 0.069       | 0.001  | <0.001     |        |
| Stage                          |     |           |            |             |        |            |        |
| I–II                           | 47  | 0.065     | 2.868      | 3.605       | 3.100  | 2.600      | 0.600  |
| III–IV                         | 57  | 0.095     | 4.028      | 4.112       | 4.112  | 4.000      | 1.400  |
| P-values*                      | 0.750| 0.02      | 0.961      | 0.150       | 0.015  | <0.001     | 0.020  |
| T                              |     |           |            |             |        |            |        |
| T1                             | 49  | 0.095     | 2.497      | 3.605       | 3.100  | 2.700      | 0.700  |
| T2–4                           | 55  | 0.081     | 4.567      | 3.784       | 5.450  | 3.700      | 1.000  |
| P-values*                      | 0.390| 0.028     | 0.787      | 0.014       | <0.001 | <0.001     | 0.011  |
| Lymphatic metastasis           |     |           |            |             |        |            |        |
| Positive                       | 59  | 0.105     | 3.784      | 3.758       | 5.450  | 4.000      | 1.400  |
| Negative                       | 45  | 0.065     | 3.078      | 4.028       | 3.100  | 2.600      | 0.700  |
| P-values*                      | 0.045| 0.381     | 0.54       | <0.001      | <0.001 | <0.001     | 0.011  |
| Distal metastasis              |     |           |            |             |        |            |        |
| Positive                       | 22  | 0.112     | 3.824      | 3.720       | 8.475  | 3.050      | 0.900  |
| Negative                       | 82  | 0.080     | 3.160      | 3.824       | 4.340  | 3.200      | 0.900  |
| P-values*                      | 0.045| 0.381     | 0.54       | <0.001      | <0.001 | <0.001     | 0.011  |

Notes: *Median of relative expression, with 25th–75th percentiles. †Mann–Whitney U test between two groups.

Abbreviations: CEA, carcinoembryonic antigen; CYFRA21-1, cytokeratin 19 fragment 21-1; LUAD, lung adenoarcinoma; LUSC, lung squamous cell carcinoma; NSCLC, non-small-cell lung cancer; SCC, squamous cell carcinoma-related antigen.
with optimal cutoff value (0.030) had sensitivity of 78.85%, specificity of 58.0%, and the corresponding AUC was 0.706 (Figure 2A). With an optimal cutoff value of 2.190, the sensitivity, specificity, and AUC of miR-21-5p were 68.27%, 78.00%, and 0.765, respectively (Figure 2B). The ROC analysis suggested the miR-223-3p could distinguish cancer patients from cancer-free subjects with sensitivity of 76.92%, specificity of 80.00%, and AUC of 0.744 (Figure 2C).
To attain better diagnostic power, we designed three panels: a microRNA panel (miR-9-5p+miR-21-5p+miR-223-3p), a tumor marker panel (CEA+CYFRA21-1+SCC), and a microRNA + tumor marker panel (miR-9-5p+miR-21-5p+miR-223-3p+CEA+CYFRA21-1+SCC). The use of the three serum microRNAs in combination (miR-9-5p+miR-21-5p+miR-223-3p) generated a sensitivity of 76.92%, specificity of 84.00%, and AUC of 0.824 (Figure 2G). The combination of CEA, CYFRA21-1, and SCC produced higher diagnostic performance, with AUC of 0.846, sensitivity of 55.77%, and specificity of 92.00% (Figure 2E). When three microRNAs and three tumor markers were combined (miR-9-5p+miR-21-5p+miR-223-3p+CEA+CYFRA21-1+SCC), this combination revealed a sensitivity of 82.69%, specificity of 88.00%, and AUC of 0.886. The microRNA + tumor marker panel yielded the best YI, +LR, and −LR scores (Figure 2F; Tables 3 and 4).

### Table 3
Area under the ROC curve analysis of serum miR-9, 21, 223, CEA, CYFRA21-1 and SCC and their combinations in differentiating NSCLC and cancer-free controls

| Parameter       | Area (95% confidence) | Asymptotic significance | Asymptotic 95% confidence intervals |
|-----------------|-----------------------|--------------------------|-------------------------------------|
| miR-9-5p        | 0.706 (0.6461–0.7766) | <0.001                   | 0.627–0.776                         |
| miR-21-5p       | 0.765 (0.705–0.822)   | <0.001                   | 0.69–0.83                           |
| miR-223-3p      | 0.744 (0.643–0.844)   | <0.001                   | 0.668–0.811                         |
| CEA             | 0.749 (0.638–0.86)    | <0.001                   | 0.673–0.816                         |
| CYFRA21-1       | 0.735 (0.649–0.832)   | <0.001                   | 0.658–0.803                         |
| SCC             | 0.616 (0.549–0.683)   | 0.019                    | 0.534–0.693                         |
| Combined miRNAs | 0.824 (0.734–0.914)   | <0.001                   | 0.754–0.88                          |
| Combined tumor markers | 0.807 | 0.0342 | <0.001 | 0.736–0.866 |
| Combined miRNAs and tumor markers | 0.886 | 0.0269 | <0.001 | 0.825–0.932 |

Notes: Combined miRNAs = miR-9-5p+miR-21-5p+miR-223-3p. Combined tumor markers = CEA+CYFRA21-1+SCC. Combined miRNAs and tumor markers = miR-9-5p+miR-21-5p+miR-223-3p+CEA+CYFRA21-1+SCC.

Abbreviations: CEA, carcinoembryonic antigen; CYFRA21-1, cytokeratin 19 fragment 21-1; NSCLC, non-small-cell lung cancer; ROC, receiver operating characteristic; SCC, squamous cell carcinoma-related antigen.

### Table 4
Diagnostic index of miR-9, 21, 223, CEA, CYFRA21-1 and SCC and their combinations based on ROC analysis results

| Parameter       | Cutoff value (area) | Sen (%) | Spe (%) | +LR | −LR | YI |
|-----------------|--------------------|--------|--------|-----|-----|----|
| miR-9-5p        | 0.030              | 87.805 | 58.000 | 1.880 | 0.360 | 0.368 |
| miR-21-5p       | 2.190              | 68.270 | 78.000 | 3.100 | 0.410 | 0.463 |
| miR-223-3p      | 2.234              | 76.920 | 80.000 | 3.850 | 0.290 | 0.569 |
| CEA             | 4.210              | 54.810 | 96.000 | 13.700 | 0.470 | 0.508 |
| CYFRA21-1       | 2.700              | 64.420 | 70.000 | 2.150 | 0.150 | 0.344 |
| SCC             | 1.400              | 55.770 | 82.000 | 1.450 | 0.700 | 0.276 |
| Combined miRNAs | –                  | 76.920 | 84.000 | 4.810 | 0.270 | 0.609 |
| Combined tumor markers | – | 55.770 | 92.000 | 6.970 | 0.480 | 0.478 |
| Combined miRNAs and tumor markers | – | 82.690 | 88.000 | 6.980 | 0.200 | 0.707 |

Notes: Combined miRNAs = miR-9-5p+miR-21-5p+miR-223-3p. Combined tumor markers = CEA+CYFRA21-1+SCC. Combined miRNAs and tumor markers = miR-9-5p+miR-21-5p+miR-223-3p+CEA+CYFRA21-1+SCC.

Abbreviations: CEA, carcinoembryonic antigen; CYFRA21-1, cytokeratin 19 fragment 21-1; LR, likelihood ratio; NSCLC, non-small-cell lung cancer; ROC, receiver operating characteristic; SCC, squamous cell carcinoma-related antigen; Sen, sensitivity; Spe, specificity; YI, Youden's index.

### Subgroup analysis
Two subgroups analyses according to region (Xuanwei and non-Xuanwei) and stage (early stage [I–II] and late stage [III–IV]) were conducted to comprehensively investigate the diagnostic role of microRNAs and tumor markers in different subgroups. The main results are summarized in Table 5.

### Table 5
Diagnostic efficiency of serum miR-9-5p, 21-5p, 223-3p, CEA, CYFRA21-1 and SCC and their combinations in subgroups by region (Xuanwei and non-Xuanwei) and stage (I–II and III–IV)

| Parameter       | Region (Xuanwei, Non-Xuanwei) | Stage (I–II, III–IV) | Sen (%) | Spe (%) | +LR | −LR | YI |
|-----------------|-------------------------------|----------------------|---------|---------|-----|-----|----|
| miR-9-5p        |                               | I–II                 | 0.699   | 0.721   |     |     |    |
| miR-21-5p       |                               | I–II                 | 0.609–0.778 | 0.611–0.814 |     |     |    |
| miR-223-3p      |                               | I–II                 | 0.732   | 0.84    |     |     |    |
| CEA             |                               | I–II                 | 0.742–0.912 | 0.742–0.912 |     |     |    |
| CYFRA21-1       |                               | I–II                 | 0.742–0.912 | 0.742–0.912 |     |     |    |
| SCC             |                               | I–II                 | 0.742–0.912 | 0.742–0.912 |     |     |    |
| Combined miRNAs |                               | I–II                 | 0.653–0.816 | 0.645–0.841 |     |     |    |
| Combined tumor markers |                               | I–II                 | 0.774   | 0.694   |     |     |    |
| Combined miRNAs and tumor markers |                               | I–II                 | 0.689–0.845 | 0.583–0.791 |     |     |    |
| Notes: Combined miRNAs = miR-9-5p+miR-21-5p+miR-223-3p. Combined tumor markers = CEA+CYFRA21-1+SCC. Combined miRNAs and tumor markers = miR-9-5p+miR-21-5p+miR-223-3p+CEA+CYFRA21-1+SCC.

Abbreviations: CEA, carcinoembryonic antigen; CYFRA21-1, cytokeratin 19 fragment 21-1; SCC, squamous cell carcinoma-related antigen.
Diagnostic role of miRNAs and tumor markers in Xuanwei and non-Xuanwei subjects
The AUCs of miR-9-5p, miR-21-5p, miR-223-3p, CEA, CYFRA21-1, and SCC in Xuanwei subjects were 0.644, 0.718, 0.752, 0.790, 0.731, and 0.583, respectively. For non-Xuanwei subjects, they were 0.730, 0.796, 0.737, 0.715, 0.745, and 0.636, respectively. The AUCs of combined miRNAs (miR-9-5p+miR-21-5p+miR-223-3p), combined tumor markers (CEA+CYPRA21-1+SCC), and combined miRNAs and tumor markers (miR-9-5p+miR-21-5p+miR-223-3p+CEA+CYPRA21-1+SCC) were 0.828, 0.855, and 0.925 in Xuanwei subjects, and 0.810, 0.798, and 0.881 in non-Xuanwei subjects, respectively. The analysis suggested these miRNAs, tumor markers, and their combination could provide reasonable diagnostic role in both Xuanwei and non-Xuanwei subjects (Table 5).

Diagnostic role of miRNAs and tumor markers in early-stage (I–II) and late-stage (III–IV) subjects
For early-stage NSCLC patients, the AUCs of miR-9-5p, miR-21-5p, miR-223-3p, CEA, CYFRA21-1, and SCC were 0.699, 0.732, 0.740, 0.774, 0.687, and 0.587, respectively. The AUCs of combined miRNAs (miR-9-5p+miR-21-5p+miR-223-3p), combined tumor markers (CEA+CYPRA21-1+SCC), and combined miRNAs and tumor markers (miR-9-5p+miR-21-5p+miR-223-3p+CEA+CYPRA21-1+SCC) were 0.811, 0.802, and 0.863, respectively. The miR-9-5p, miR-21-5p, miR-223-3p, CEA, CYFRA21-1, and SCC could discriminate between advanced NSCLC patients and cancer-free subjects with AUCs of 0.721, 0.840, 0.752, 0.694, 0.843, and 0.840, respectively. For combined miRNAs (miR-9-5p+miR-21-5p+miR-223-3p), combined tumor markers (CEA+CYPRA21-1+SCC), and combined miRNAs and tumor markers (miR-9-5p+miR-21-5p+miR-223-3p+CEA+CYPRA21-1+SCC), the AUCs were 0.858, 0.873, and 0.945, respectively. Generally, the diagnostic performance of miRNAs and tumor markers were higher in patients with advanced NSCLC than in early-stage patients (Table 5).

Discussion
In the present study, we investigated expression levels of serum miR-9-5p, miR-21-5p, miR-223-3p, miR-135b-5p, miR-339-5p, and miR-501-5p in NSCLC patients and cancer-free subjects. We found that only miR-9-5p, miR-21-5p, and miR-223-3p were highly expressed in NSCLC patients when compared with cancer-free subjects. Further analysis suggested these three miRNAs could distinguish NSCLC patients from cancer-free subjects with AUCs that ranged from 0.706 to 0.765. The AUCs of miR-9-5p, miR-21-5p, and miR-223-3p were 0.706, 0.765, and 0.744, respectively, suggesting these three miRNAs could represent potential biomarkers for the diagnosis of NSCLC in Yunnan residents. The panel of three miRNAs could differentiate lung cancer patients from cancer-free subjects with an AUC of 0.824. When the panel of three miRNAs was combined with CEA, CYFRA-21, and SCC, this combination could yield a sensitivity of 82.69%, a specificity of 88.00%, and an AUC of 0.866. In both Xuanwei and non-Xuanwei subjects, these miRNAs, especially when combined with tumor markers, could play a reasonable diagnostic role.

As an oncomir, upregulation of miR-21-5p leads to tumor development and progression.26,27 Circulating miR-21-5p could serve as an ideal diagnostic and prognostic biomarker.26,27 Our analysis suggested serum miR-21-5p could distinguish NSCLC patients from cancer-free subjects, with an AUC of 0.765 in the Yunnan populace. Furthermore, its diagnostic performance in Xuanwei subjects was moderate (with an AUC of 0.718).

The expression of miR-9-5p was increased in human NSCL tissues when compared with adjacent noncancerous tissues.32 Upregulated miR-9-5p expression correlated with adverse clinical features and unfavorable survival.23 Fei et al reported serum miR-9-5p was significantly upregulated in patients with osteosarcoma when compared with healthy controls.24 However, a study by Sun et al found that the expression level of serum miR-9-5p was significantly downregulated in patients with oral squamous cell carcinoma when compared with healthy controls.28 Sromek et al found there was no difference in plasma miR-9-5p expression between NSCLC patients and healthy controls. However, the plasma miR-9-5p expression level would decrease after 1 month of surgery.30 The expression levels of circulating miR-9-5p were conflicting in various cancers;28-30 and only one study investigated the diagnostic role of miR-9-5p in lung cancer.30 In our study, we found serum miR-9-5p was upregulated in NSCLC patients when compared with cancer-free subjects. The AUC of serum miR-9-5p was 0.706. To determine and validate whether serum miR-9-5p could serve as a useful diagnostic biomarker for NSCLC, a larger sample size would be needed.

Cell-line experiments suggested miR-223-3p was upregulated in human lung cancer A549 cells, and it promoted tumor progression via activation of the NF-xB signaling pathway.31 Circulating miR-223-3p was also reported to be upregulated...
in NSCLC patients when compared with controls and could serve as a diagnostic biomarker in lung cancer patients.\textsuperscript{32-34} Our results showed the AUC of serum miR-223-3p was 0.744. In Xuanwei subjects, the AUC was 0.752. It was lower than the AUC in the study reported by Zhang et al (AUC =0.809)\textsuperscript{32} and Geng et al (AUC =0.94).\textsuperscript{34}

### Racial disparities and selection bias

We did not find significant differences in miR-135b-5p, 339-5p, and 501-5p between NSCLC patients and cancer-free subjects. miR-135b-5p was upregulated in some cancers such as gastric cancer,\textsuperscript{35} colorectal cancer,\textsuperscript{36} and breast cancer.\textsuperscript{37} MiR-135b-5p was upregulated in highly invasive NSCLC cells; expression of miR-135b-5p enhanced cancer cell invasive and migratory abilities and promoted cancer metastasis by regulating multiple targets in the Hippo pathway and LZTS1.\textsuperscript{16} MiR-339-5p could promote colony formation and attenuate apoptosis of lung carcinoma cell lines through targeting α1,2-fucosyltransferase 1 and regulation of the downstream protein Lewis y.\textsuperscript{38} MiR-501-5p was upregulated in hepatocellular carcinoma specimens and cell lines. However, its potential role in lung cancer and other cancers was seldom investigated. Our previous studies suggested miR-501-5p was upregulated in NSCLC tissues.\textsuperscript{39} All three miRNAs (miR-135b-5p, 339-5p, and 501-5p) in the serum showed no differences of expression between NSCLC patients and cancer-free subjects. Therefore, the circulating miRNA profile may not be consistent with the tissue miRNA profile.

CEA, CYFRA21-1, and SCC are widely used as serum diagnostic biomarkers for NSCLC. CEA is a cell-adhesion molecule associated with adverse clinical outcomes for certain cancers by promotion of invasion, dissemination, metastasis, and immune suppression.\textsuperscript{40} Serum CEA is a useful, supportive, diagnostic tool for NSCLC. More recent studies found that higher serum CEA levels were more frequently detected in LUAD than LUSC.\textsuperscript{41,42} We only found the trend that the CEA level was higher in LUAD than in LUSC, although the \( P \)-value was not significant (4.55 vs 3.18, \( P=0.069 \)). The limited sample size might be one reason; if more samples were included, the result may have become significant. CYFRA21-1 is a keratin intermediate filament protein belonging to the type I keratin family.\textsuperscript{43} An elevated CYFRA21-1 level was associated with advanced stage, mediastinal lymph node metastasis, and unresectable tumors.\textsuperscript{9} SCC is a cytoskeletal protein fraction. The sensitivity of SCC was lower than CEA (35% vs 56%), but the specificity of SCC was higher than that of CEA.\textsuperscript{9} SCC might be more useful for LUSC; however, both CEA and CYFRA21-1 were suitable markers for LUAD and LUSC.\textsuperscript{9} This could possibly be one of the reasons that the SCC showed a lower AUC than CEA or CYFRA21-1 (0.616 vs 0.749 or 0.735).

Single miRNA or tumor marker had limited diagnostic power. Their AUC ranged from 0.616 to 0765. Among these biomarkers, miR-223-3p and CEA had the best YI, +LR, and -LR scores, which indicated that they were suitable indicators for the diagnosis of NSCLC. Combined miRNAs and tumor markers could yield satisfactory diagnostic performance. A subgroup analysis according to the region (Xuanwei vs non-Xuanwei) was conducted. We found miR-223-3p (AUC =0.752) and CEA (AUC =0.791) might be more suitable in discriminating NSCLC from cancer-free subjects in Xuanwei. However, in the non-Xuanwei population, miR-21-5p and CYFRA21-1 had higher AUCs. However, this conclusion should be interpreted with caution as the sample size was limited.

Subgroup analysis by stage (I–II vs III–IV) suggested that both miRNAs and tumor markers had lower AUCs in stages I–II NSCLC when compared with stages III–IV disease. Possibly, the greater the tumor load in the body, the higher the secretion of miRNAs and tumor markers from living cancer cells, or they may be released from degraded tumor cells into the circulation. Combining both miRNAs and tumor markers together (miR-9-5p+miR-21-5p+miR-223-3p+CEA+CYFRA21-1+SCC) could yield satisfactory diagnostic power to identify NSCLC at an early stage, which may be greatly beneficial to improve the outcome and prognosis of the disease.

Our study has some limitations. Firstly, the sample size was not very large and all patients were from the same center. To further confirm the diagnostic role of these miRNAs in NSCLC patients in Yunnan and in Xuanwei, future multicenter, randomized, prospective trials in a larger patient population are needed.

Secondly, we only included cancer-free subjects as controls. To comprehensively investigate the diagnostic role of miRNAs in the screening or differential diagnosis of lung cancer, healthy control and cancer-free controls (including benign pulmonary nodules as well as other benign pulmonary disease such as pneumonia and chronic obstructive pulmonary disease) should be included.

Finally, we selected miRNAs according to microarray analysis results from tissue specimens and reports in the literature. This approach could not reflect the real miRNA profile in serum. Identifying a potential miRNA profile by microarray analysis or next-generation sequencing technology in serum might be more reasonable.
In conclusion, our analysis suggested miR-9-5p, miR-21-5p, and miR-223-3p could serve as potential diagnostic biomarkers in NSCLC in Yunnan. To achieve satisfactory diagnostic performance, a combination of miRNAs and tumor markers is recommended. Given the limitation in our study, further multicenter, randomized, prospective trials in a larger population are needed in order to confirm the results. Furthermore, more innovative, noninvasive, sensitive, and reliable biomarkers still need to be discovered to improve the prognosis of NSCLC.

Acknowledgments
This work was supported by the PHD Innovation Fund of Kunming Medical University (number 2017D005), the Project of Basic Applied Research in Yunnan Province (number 2016FB145), the National Natural Science Foundation of China (number 81460441), the Internal Institution Project of Ministry of Health of Yunnan Province (number 2017NS200), and the Health Science and Technology Project of Yunnan Province (number 2014NS001).

Disclosure
The authors report no conflicts of interest in this work.

References
1. Torre LA, Bray F, Siegel RL, Ferlay J, Lortet-Tieulent J, Jemal A. Global cancer statistics, 2012. CA Cancer J Clin. 2015;65(2):87–108.
2. Chen W, Zheng R, Baade PD, et al. Cancer statistics in China, 2015. CA Cancer J Clin. 2016;66(2):115–132.
3. He XZ, Chen W, Liu ZY, Chapman RS. An epidemiological study of lung cancer in Xuan Wei County, China: current progress. Case-control study on lung cancer and cooking fuel. Environ Health Perspect. 1991;94:9–13.
4. Chen G, Sun X, Ren H, et al. The mortality patterns of lung cancer between 1990 and 2013 in Xuanwei, China. Lung Cancer. 2015;90(2):155–160.
5. Chen Z. The Third National Retrospective Surveys on Causes of Death. Beijing: Peking Union Medical College Press; 2008.
6. Li GJ, Huang YC, Tian LW, et al. [Relationship between high incidence of lung cancer among non-smoking women and silica in C1 bituminous coal in Xuanwei, Yunnan Province, China]. Zhonghua Lao Dong Wei Sheng Za Zhi. 2013;31(1):30–36. Chinese.
7. Mumford JL, He XZ, Chapman RS, et al. Lung cancer and indoor air pollution in Xuan Wei, China. Science. 1987;235(4785):217–220.
8. Luo Q, He X, Shen M, et al. Variation in lung cancer risk by smoky coal subtype in Xuanwei, China. Int J Cancer. 2008;123(9):2164–2169.
9. Nakamura H, Nishimura M. History, molecular features, and clinical importance of conventional serum biomarkers in lung cancer. Surg Today. 2017;47(9):1037–1059.
10. Kong YW, Ferland-McCollough D, Jackson TJ, Bushell M. microRNAs in cancer management. Lancet Oncol. 2012;13(6):e249–e258.
11. Chen X, Liang H, Zhang J, Zen K, Zhang CY. Secreted microRNAs: a new form of intercellular communication. Trends Cell Biol. 2012;22(3):125–132.
12. Xu T, Liu X, Han L, Shen H, Liu L, Shu Y. Up-regulation of miR-9 expression as a poor prognostic biomarker in patients with non-small cell lung cancer. Clin Transl Oncol. 2014;16(5):469–475.
36. He Y, Wang J, Wang J, et al. MicroRNA-135b regulates apoptosis and chemoresistance in colorectal cancer by targeting large tumor suppressor kinase 2. *Am J Cancer Res*. 2015;5(4):1382–1395. eCollection 2015.

37. Hua K, Jin J, Zhao J, et al. miR-135b, upregulated in breast cancer, promotes cell growth and disrupts the cell cycle by regulating LATS2. *Int J Oncol*. 2016;48(5):1997–2006.

38. Gan CZ, Li G, Luo QS, Li HM. miR-339-5p downregulation contributes to Taxol resistance in small-cell lung cancer by targeting α1,2-fucosyltransferase 1. *IUBMB Life*. 2017;69(11):841–849.

39. Chen S, Zhou YC, Chen Y, et al. [Expression profile of miR-501-5p in lung adenocarcinoma patients from Xuanwei area]. *Nan Fang Yi Ke Da Xue Xue Bao*. 2017;37(3):354–359. Chinese.

40. Beauchemin N, Arabzadeh A. Carcinoembryonic antigen-related cell adhesion molecules (CEACAMs) in cancer progression and metastasis. *Cancer Metastasis Rev*. 2013;32(3–4):643–671.

41. Molina R, Filella X, Augé JM, et al. Tumor markers (CEA, CA 125, CYFRA 21-1, SCC and NSE) in patients with non-small cell lung cancer as an aid in histological diagnosis and prognosis. Comparison with the main clinical and pathological prognostic factors. *Tumour Biol.* 2003;24(4):209–218.

42. Pollán M, Varela G, Torres A, et al. Clinical value of p53, c-erbB-2, CEA and CA125 regarding relapse, metastasis and death in resectable non-small cell lung cancer. *Int J Cancer*. 2003;107(5):781–790.

43. Fuchs E, Weber K. Intermediate filaments: structure, dynamics, function, and disease. *Annu Rev Biochem*. 1994;63:345–382.