International Journal of Nanomedicine

Dovepress
Taylor & Francis Group

Open Access Full Text Article

ORIGINAL RESEARCH

Plasma Small Extracellular Vesicle microRNAs as Non-Invasive Biomarkers for Lung Cancer Detection

Qiaoli Lv[1,*], Yangzhong Guo[1,*], Xiaoya Xu[2,*], Dongyu Liu[2], Xiaoling Xiong[1], Qingfeng Wei[1], Yan Feng[2], Dadong Zhang[2], Zhisheng He[1], Weimin Mao[1]

[1]Thoracic Oncology Laboratory, Jiangxi Key Laboratory of Oncology, Jiangxi Cancer Hospital & Institute, The Second Affiliated Hospital of Nanchang Medical College, Nanchang, Jiangxi, 330029, People's Republic of China; [2]Department of Clinical and Translational Medicine, 3D Medicines Inc., Shanghai, 201114, People's Republic of China

*These authors contributed equally to this work

Correspondence: Weimin Mao; Zhisheng He, Thoracic Oncology Laboratory, Jiangxi Key Laboratory of Oncology, Jiangxi Cancer Hospital & Institute, 519 Beijing East Road, Nanchang, Jiangxi, 330029, People's Republic of China, Email maowm@zjcc.org.cn; hzs1973@sina.com

International Journal of Nanomedicine downloaded from https://www.dovepress.com/
For personal use only.

Background: Current non-invasive approaches for lung cancer (LC) detection exhibit inherent limitations in diagnostic accuracy, or inadequate clinical validation. Consequently, there exists an urgent unmet need for rigorously validated, non-invasive biomarkers exhibiting high sensitivity and specificity to enable the early detection of LC.

Methods: We employed small RNA sequencing technology to detect microRNA (miRNA) expression in small extracellular vesicle (sEV) isolated from plasma samples of study participants. The collected samples were subjected to retrospective analysis. A diagnostic model was developed (n = 80) and validated (n = 52) to discriminate between non-malignant controls (NCs, comprising healthy individuals and benign lesions cases) and patients with LC (Stages I/II). Model performance was rigorously evaluated using several metrics, with the area under the curve (AUC) serving as the primary metric.

Results: The small RNA sequencing analysis of plasma sEV miRNA identified distinct expression signatures (14 differentially expressed sEV miRNAs) between NCs and LC samples. The diagnostic model with the best performance was constructed using hsa-miR-423-5p, hsa-miR-340-3p, hsa-miR-320b, hsa-miR-98-5p, hsa-miR-26a-5p, hsa-miR-193b-5p, hsa-miR-629-5p, and hsa-miR-92b-5p. The diagnostic model achieved an AUC of 0.956, a sensitivity of 94%, and a specificity of 93% in the training cohort and an AUC of 0.985, a sensitivity of 86%, and a specificity of 97% in the validation cohort.

Conclusion: Our findings demonstrates that plasma sEV miRNA exhibits a highly discriminative biomarker for distinguishing NCs group from early malignant lesions, making it a promising tool for auxiliary detection of early-stage LC.

Keywords: lung cancer, early-stage, non-invasive biomarkers, small RNA sequencing, small extracellular vesicle miRNA

Introduction

Lung cancer (LC) remains a predominant global health crisis, ranking as the leading cause of cancer-related mortality worldwide. According to global cancer statistics, it is estimated that over 2 million people are newly diagnosed with LC each year, making it one of the cancers that seriously threaten human health.[1–3] Histologically, LC is categorized into two primary subtypes: non-small cell lung cancer (NSCLC), comprising 80–85% of diagnoses, and small cell lung cancer (SCLC). NSCLC encompasses diverse pathological entities, including lung adenocarcinoma (LUAD), squamous cell carcinoma (SCC), and large cell undifferentiated carcinoma,[4] each exhibiting distinct molecular profiles and clinical behaviors. Due to the concealed clinical symptoms and the lack of effective screening methods for early clinical diagnosis, approximately 75% of LC patients present with locally advanced or metastatic stage at initial diagnosis, missing the optimal window for radical surgical treatment and facing a poorer prognosis.[5] This advanced presentation precludes curative surgical intervention, contributing to dismal 5-year survival rates (≤20% for stage IV) compared to 77–92% observed in early-stage (stage I) patients undergoing resection.[6] Therefore, detection of early-stage LC is crucial for improving patient prognosis.

Received: 14 April 2025
Accepted: 26 August 2025
Published: 8 September 2025

© 2025 Lv et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at https://www.dovepress.com/terms.php and incorporate the Creative Commons Attribution – Non Commercial (unported, v4.0) License (http://creativecommons.org/licenses/by-nc/4.0/). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (https://www.dovepress.com/terms.php).

Bronchoscopy and histopathological biopsy are regarded as the gold standards for diagnosing LC. However, bronchoscopy exhibits variable sensitivity in diagnosing early-stage lesions, while histopathological biopsy, being an invasive procedure, poses risks of infection, bleeding, and tumor dissemination.[7–9] Liquid biopsy, with its merits of non-invasive sampling and the possibility of multiple sampling, has been utilized in early diagnosis, screening, postoperative monitoring, and treatment decision-making for LC.[10–12] Among them, sputum cytology serves as an adjunctive diagnostic tool for LC, which is non-invasive and convenient but has low sensitivity. Commonly used blood-based tumor markers such as carcinoembryonic antigen (CEA), cytokeratin 19 fragment (CYFRA21-1), neuron-specific enolase (NSE), squamous cell carcinoma antigen (SCC-Ag), and pro-gastrin-releasing peptide (proGRP) exhibit limited sensitivity and specificity in the adjunctive diagnosis of LC and are prone to false positives.[7–10,13–15] Other blood-based diagnostic biomarkers for LC, such as miRNAs, ctDNA, and cfDNA methylation, either struggle to balance sensitivity and specificity, a lack of validation, or have been studied with small sample sizes, rendering them ineffective as differential diagnostic biomarkers.[9–12,15–19]

Small extracellular vesicle (sEV) are membrane-bound vesicles with a lipid bilayer that are secreted and released by cells. sEV are rich in content and exhibit strong stability, which gives them an advantage in liquid biopsy, particularly in the early diagnosis of tumors, where they hold great potential.[20,21] It has been reported that sEV miRNAs, sEV mRNAs, sEV lncRNAs, and sEV proteins have diagnostic applications in LC. MiRNAs are highly enriched and stable in sEV, extensively participating in various mechanisms that regulate LC proliferation and invasion. Consequently, sEV miRNAs demonstrate greater potential and novelty in the diagnosis of LC.[10,15,22–27] Among these, diagnostic models with superior performance have exhibited sensitivities and specificities that exceed 90%.[23,27,28] However, these studies have several limitations. Firstly, they lack independent validation cohorts. Secondly, they involve relatively small sample sizes. It is therefore necessary to discover and validate biomarkers that can enhance the specificity of early-stage LC diagnosis.

In this study, we included a total of 132 blood samples from healthy individuals, patients with benign lesions, and patients with LC (Stages I/II). Plasma sEV were extracted, and small RNA sequencing was performed to detect changes in the miRNA expression profile of blood sEV. In two independent cohorts, an sEV miRNAs-based biomarker diagnostic model for LC detection was constructed and validated. Finally, pathway enrichment analysis was carried out on the sEV miRNAs involved in the diagnostic model to further explore the potential of sEV miRNAs as biomarkers for early-stage LC detection.

Materials and Methods

Patient Recruitment and Blood Sample Collection

This study enrolled a total of 132 patients who were divided into a training cohort (n = 80) and a validation cohort (n = 52). The training cohort included 30 healthy individuals and 16 patients with benign lesions and 34 patients with LC (Stages I/II), while the validation cohort comprised 20 healthy individuals and 11 patients with pulmonary benign lesions and 21 patients with LC (Stages I/II). Blood samples were acquired from all patients between July 2021 and October 2023, with specific clinical information detailed in Table 1. The collection of blood samples in this study was approved by the Ethics Committee of Jiangxi Provincial Cancer Hospital (No. 2021ky231).

Blood samples were collected using EDTA anticoagulant tubes (REF367525, BD, USA) and pretreatment was conducted within 24 hours. Initially, the blood collection tubes were centrifuged at 1600 g for 10 minutes at 4°C to evaluate the levels of hemolysis, with only samples that received a grade of 4 or lower being retained for further analysis. Subsequently, the supernatant was aspirated and transferred to 1.5 mL EP tubes and centrifuged again at 16,000 g for 15 minutes at 4°C to remove residual cellular debris. Finally, the supernatant was divided into 1.5 mL EP tubes and stored at −80°C for future use.

Isolation of Plasma sEV

Plasma sEV were isolated using a plasma sEV isolation reagent (L3525, 3DMed, Shanghai, China) developed by Shanghai 3D Medicines Inc. Frozen plasma samples were thawed at 37°C water bath, followed by centrifugation at 12,000 g for 10 minutes at 4°C. The supernatant was then filtered sequentially through a 0.45 μm filter column (CLS8163-100EA, Corning, USA) and a 0.22 μm filter column (CLS8161-100EA, Corning, USA) at 12,000 g for

Table 1 The Demographic and Clinicopathologic Characteristics of Participants

Characteristics	Training Cohort (n = 80)	Validation Cohort (n = 52)
Categories, n (%)		
Healthy	30 (37.5)	20 (38.5)
Benign	16 (20.0)	11 (21.2)
Malignant	34 (42.5)	21 (40.4)
Age, mean (SD, Benign and Malignant)	60.2 (8.1)	62.9 (7.9)
Gender, n (%, Benign and Malignant)		
Female	21 (42.0)	17 (53.1)
Male	29 (58.0)	15 (46.9)
Benign, n (%)		
Semi-positive Nodule	5 (31.2)	4 (36.4)
Positive Nodule	6 (37.5)	3 (27.3)
Other	5 (31.2)	4 (36.4)
Malignant stages, n (%)		
I	28 (82.4)	17 (81.0)
II	6 (17.6)	4 (19.0)
Malignant Subgroup, n (%)		
Lung Squamous Cell Carcinoma	8 (23.5)	4 (19.0)
Lung Adenocarcinoma	26 (76.5)	16 (76.2)
Small Cell Lung Cancer	0 (0.0)	1 (4.8)

5 minutes at 4°C. After measuring filtered volumes, 0.25 times the plasma sample volume of L3525 reagent were added, mixed thoroughly, and incubated at 4°C for 30 min. The mixture was centrifuged at 4700 g for 30 min at 4°C, and sEV pelleted were re-suspended in 200 μL phosphate-buffered saline (PBS, pH 7.4).

TEM Identification of Plasma sEV

Plasma sEV were first fixed with 4% paraformaldehyde to examine their morphology using transmission electron microscopy (TEM). After fixation, the samples were transferred to a carbon-coated copper grid for electron microscopy. The grid was washed with PBS, PBS containing 50 mM glycine, and PBS containing 0.5% BSA. Subsequently, the grid was stained with 2% uranyl acetate. A transmission electron microscope (H-7650, Hitachi High-Technologies, Japan) was used to analyze the morphological characteristics of plasma sEV.

NTA Identification of Plasma sEV

Plasma sEV diluted with PBS were analyzed using a NanoSight NS300 instrument (Malvern, UK). Using the 488 nm excitation module, the camera lens parameters were set with a shutter value of 890, a gain value of 146, and a detection threshold of 7. At least 200 complete tracks were analyzed and acquired for each video. Finally, the nanoparticle tracking analysis (NTA) software (version 2.3) was used to analyze the nanoparticle tracking data of plasma sEV.

NanoView Identification of Plasma sEV

Plasma sEV were mixed thoroughly with PBS and diluted with 1X sample buffer. Then 50 μL of the diluted plasma sEV solution was loaded onto ExoView exosome detection kit chips, and incubated at room temperature for 16 hours. Afterward, 1X buffer A was added to each well of the chip, and shaken on a horizontal shaker at 500 rpm for 3 minutes. This washing step was repeated four times. Subsequently, the prepared CD9, CD63, and CD81 staining solutions were added to each well. Shake the chip slowly on the shaker for 1h away from light, then washed and placed on absorbent paper to dry. Finally, the chip was loaded into an ExoView R100 (NanoView Biosciences, Boston, MA, USA) for detection.

Construction of the Plasma sEV-miRNA Library

Total RNA from plasma sEV was extracted using a miRNeasy Serum/Plasma Kit (217184, QIAGEN, Shanghai, China). The concentration and fragment distribution of the extracted RNA samples was measured using the Small RNA Analysis Kit (5067-1548, Agilent, Santa Clara, California, USA). Subsequently, a NEBNext Multiplex Small RNA Library Prep Set for Illumina Kit (E7300L, NEB, USA) was used for library construction, following the specific instructions provided in the product manual. The library preparation process can be summarized as follows: Each RNA sample (6 μL) sequentially underwent ligation of the 3′ adapter, hybridization of the reverse transcription primer, ligation of the 5′ adapter, reverse transcription, and PCR amplification for 18 cycles, and the products were purified using the NucleoSpin Gel and PCR Clean-up Kit (740609.250, MN, Germany). The concentration of the library was determined using a Qubit 3.0 fluorometer, and the fragment size distribution was analyzed using an Agilent 2100 Bioanalyzer with the corresponding chip and reagents (5067-4626, Agilent, Santa Clara, California, USA). Qualified libraries were sequenced using the Illumina NovaSeq6000 system with a sequencing strategy of 2 × 150-bp, generating approximately 6 G of sequencing data per library.

Small RNA Sequencing Data Processing

After the 3′-adaptors in raw FASTQ reads were trimmed with Cutadapt,[29] Reads were mapped to the human reference sequence (hg19) using BWA (0.7.12-r1039) [arXiv:1303.3997v2]. The annotation of miRNAs was performed using the Gencode (v25)[30] and miRbase (v21)[31] databases. The total number of reads mapped to the miRNA region in each sample exceeded 1.5 million nucleotides. The miRNAs that were sequenced to a depth of at least 10 were used for subsequent analysis. Trimmed mean of M-values (TMM) method was used to normalize the raw miRNA expression data from the training cohort and validation cohort.

Candidate miRNA Markers

Non-malignant controls (NCs, contain healthy and benign lesions) and patients with LC (Stages I/II) were used as diagnostic targets. Candidate diagnostic miRNA markers were screened from the training cohort using statistical methods.

Statistical differences in the expression levels of miRNAs between the two groups (NCs and LC) or the three groups (healthy, benign lesions and LC) were identified. The miRNAs that demonstrated differential expression between NCs and LC samples, while exhibiting concordance between healthy and benign lesions samples, were used as candidate markers for subsequent analysis.

In order to explore the targeted biological pathways of the miRNA markers, the DIANA-miRPath (v4.0) tool was utilized with the Kyoto Encyclopedia of Genes and Genomes (KEGG)[32] and Reactome[33] databases.

Construction of Diagnostic Model

In order to identify the final markers and construct the diagnostic model, a variety of machine learning (ML) algorithms with different hyper-parameters were used, including linear, tree-based ensemble models and neural network. The importance of the selected miRNA markers was evaluated using the permutation importance method. The model performance was evaluated using the mean of 5-fold cross-validation in the training cohort, applying different metrics, including the AUC and recall of each target.

The primary procedure was as follows: (a) The importance of each miRNA was assessed using a permutation feature importance technique, which shuffles the values of a single feature randomly and observes the resulting degradation of the model's score. miRNAs with an importance value exceeding 0 were retained, indicating a positive effect on models. (b) The performance of different machine learning models with retained candidate miRNAs were evaluated.

This process was iterated until the importance of all miRNAs retained exceeded 0. The final models were trained based on the entire training cohort.

The validation cohort was used to assess the generalization performance of the models with the corresponding markers.

Quantitative Reverse-transcription PCR

The reverse transcription of sEV miRNAs was performed using the TaqMan™ Advanced miRNA cDNA Synthesis Kit (A28007, Applied Biosystems™, Carlsbad, California, USA). The reverse transcription products were then subjected to detection of the expression levels of miR-340-3p, miR-98-5p, miR-26a-5p, miR-423-5p, miR-320b, miR-193b-5p, miR-629-5p, and miR-92b-5p using specific probes (A25576, Applied Biosystems™, Carlsbad, California, USA) and reagents (4444557, Applied Biosystems™, Carlsbad, California, USA) in conjunction with the Applied Biosystems 7500 Fast Real-Time PCR System. For detailed procedures, please refer to the product manual. Following previous reports, miR-451a was selected as the internal reference gene,[34] and the relative expression levels of sEV miRNAs were calculated using the $2^{-\Delta\Delta Ct}$ method.

Statistical Analysis and Machine Learning

The statistical tests and machine learning algorithms were conducted using the Python (v3.12.8) packages and a custom script. The Mann-Whitney *U*-test and Kruskal-Wallis *H*-test, as implemented in SciPy (v1.14.1), were employed to identify statistically significant differences in miRNA expression among groups. A p-value of less than 0.05 was considered to indicate a statistically significant difference. Scikit-learn (v1.5.2) were utilized to perform data transformation and feature selection using machine learning algorithms. AutoGluon (v1.2) was utilized to construct diagnostic model and perform hyper-parameter tuning.

Results

Study Design and Patient Clinical Information

To investigate the potential of plasma sEV miRNA as a diagnostic biomarker for early-stage LC, we collected blood samples from healthy individuals, patients with benign lesions and early malignant lung lesions, after enriching the sEV, we employed miRNA high-throughput sequencing to discover and validate sEV miRNA biomarkers for LC detection, as shown in Figure 1A.

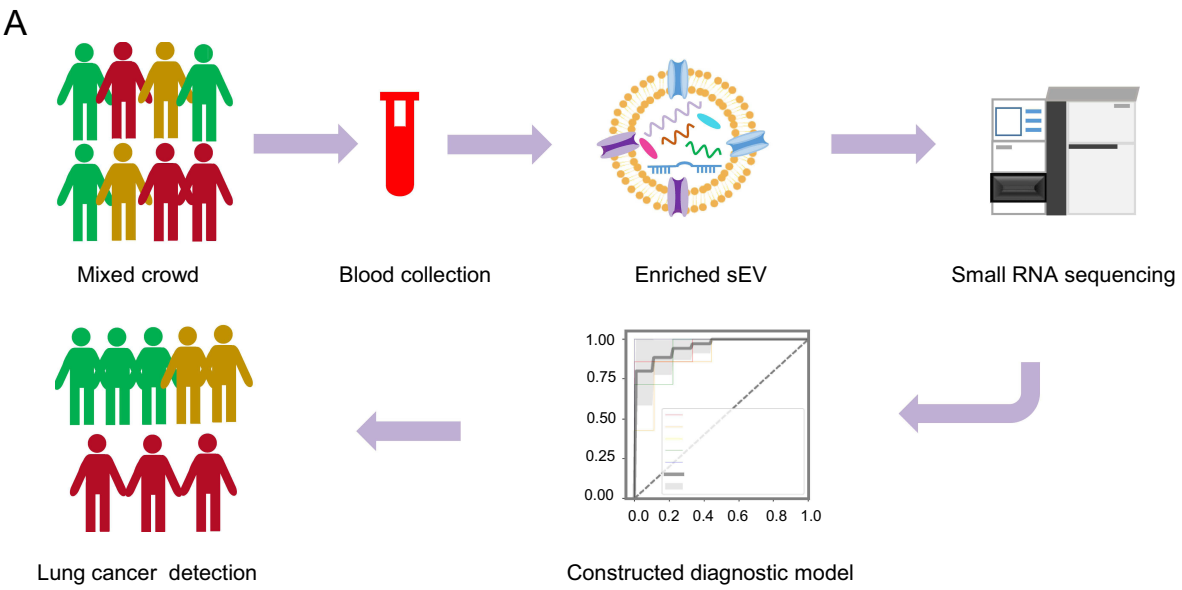
The specific workflow is illustrated in Figure 1B, we recruited a total of 150 participants, including healthy individuals, benign and early malignant lung lesions. After excluding patients who had blood samples with hemolysis, incomplete pathological information, or for which library construction failed, 132 patients were included and were divided into two cohorts. The training cohort comprised information on 30 healthy individuals, 16 benign lesions patients, and 34 patients with LC (Stages I/II), while the validation cohort included 20 healthy individuals, 11 benign lesions patients, and 21 LC samples. Table 1 presents patient age, sex, tumor stage, and smoking information. Among the LC patients, 76.4% had LUAD, and 21.8% had SCC, and 1.8% had SCLC. There were no significant differences in clinical characteristics such as age and gender between the training and validation cohorts.

Characterization of Plasma sEV

To assess the purity and yield of the extracted plasma sEV, we used TEM, NTA, and sEV particle size analysis (NanoView) to examine the morphology, particle size, and characteristic protein expression of the sEV. The main peaks of the plasma sEV particle size detected by NTA were 50 nm and 57 nm, which were consistent with the particle size distribution of sEV, as shown in Figure 2A. TEM revealed a typical cup-shaped morphology of plasma sEV, as shown in Figure 2B. Additionally, the expression of the proteins CD9, CD63, and CD81, which are markers of sEV, was detected in the plasma sEV, as shown in Figure 2C.

Construction and Validation of an Early Diagnostic Model for Differentiating Between NCs and Patients with LC Using Plasma sEV miRNA

We extracted miRNA from plasma sEV and performed small RNA sequencing to validate the potential of plasma sEV miRNA as a biomarker for differentiating between NCs and patients with LC (Stages I/II). In the training cohort, based on the screening thresholds of $p<0.05$, and there was no difference in miRNA between healthy and benign lesions, 14 differentially expressed sEV miRNAs (DEMs) were identified, as shown in Figure 3A. Furthermore, these DEMs were



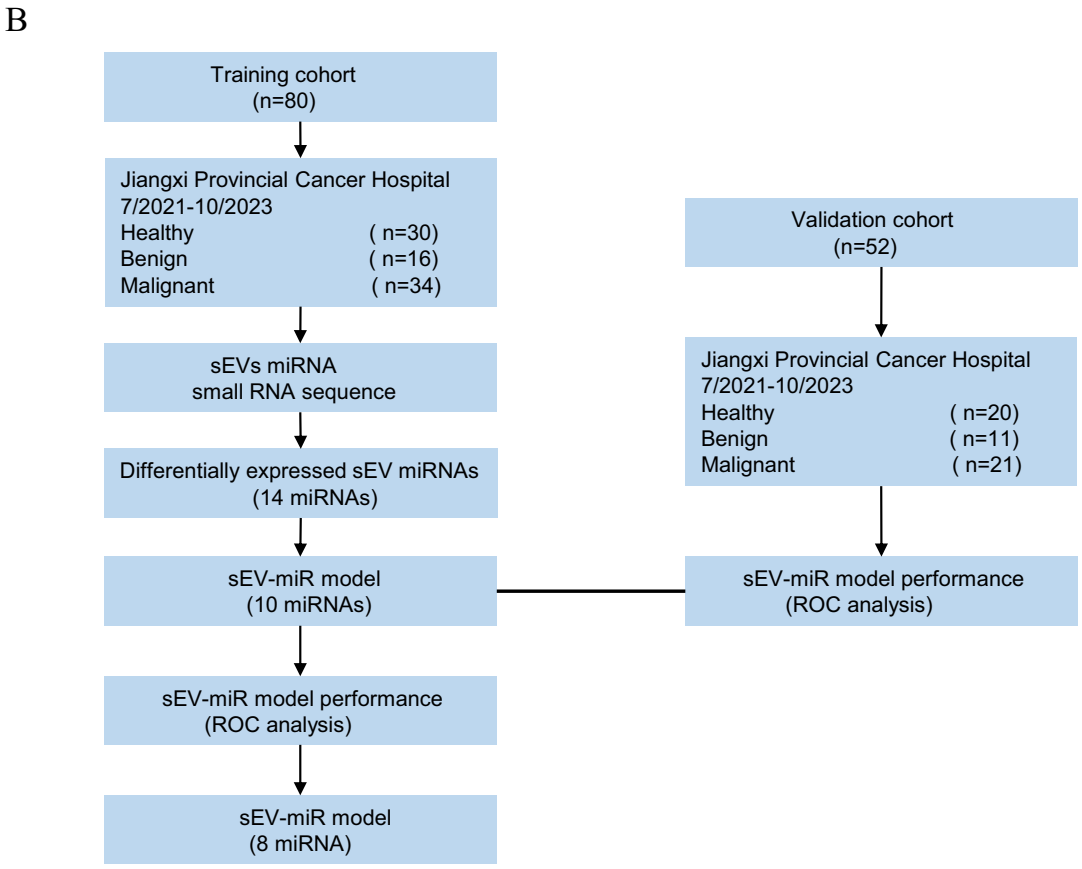


Figure 1 Study Design and Clinical Cohorts. (**A**) Is a schematic diagram illustrating the use of sEV miRNA liquid biopsy technology to differentiate between non-malignant controls (NCs) and patients with early-stage lung cancer (LC). (**B**) Presents a flowchart of the construction and validation of the diagnostic model for early-stage LC.

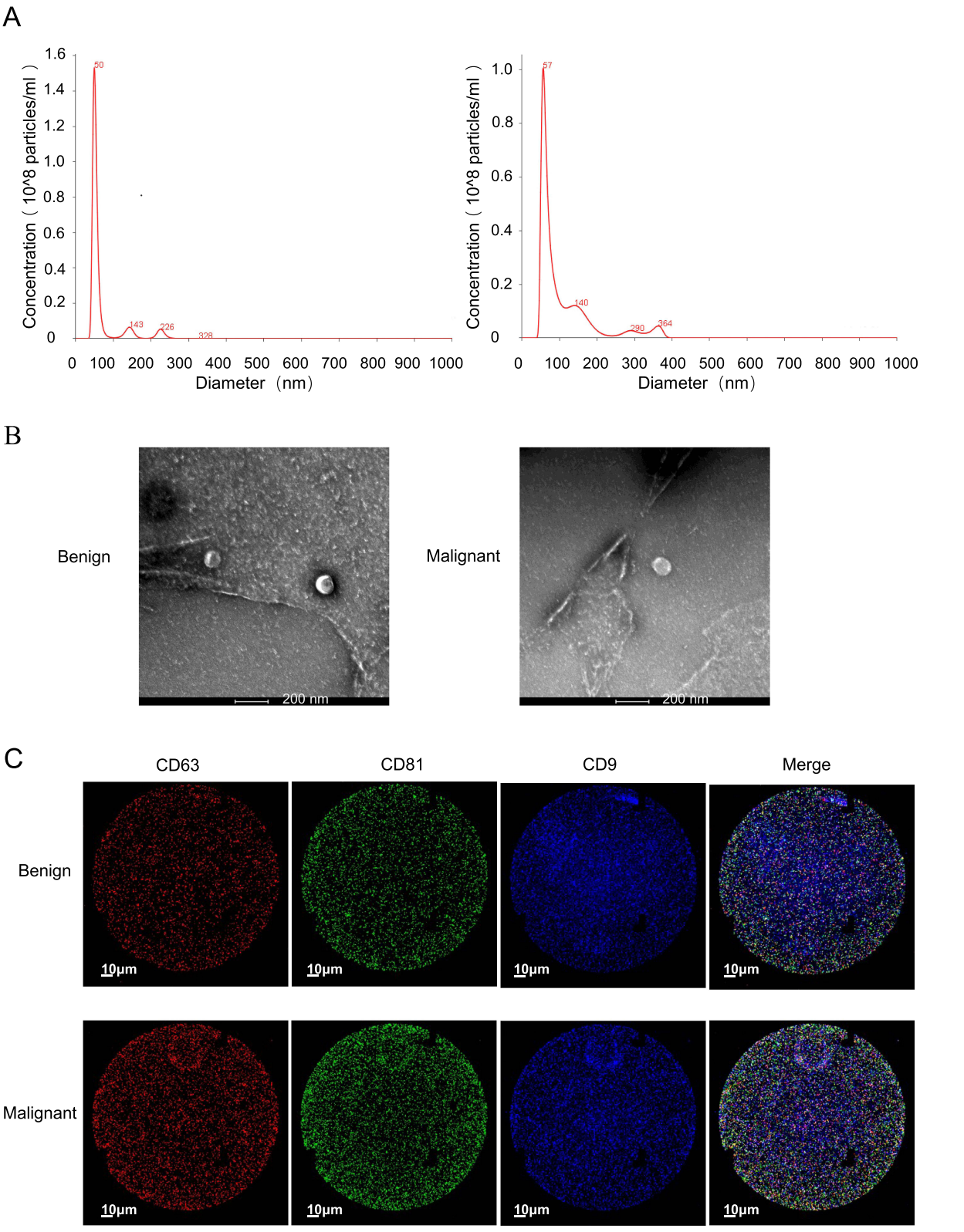


Figure 2 Characterization of sEVs in NCs and Patients with Early-Stage LC. Representative results of the particle size of sEVs were detected by NTA (**A**), the morphology of sEVs was detected by TEM (**B**), and the membrane protein of sEVs was detected by NanoView sEVs analysis (**C**) from NCs and patients with early-stage LC. The numbers in Figure 2A represent the size of the main peak for sEVs particle diameter.

https://doi.org/10.2147/IJN.S534378

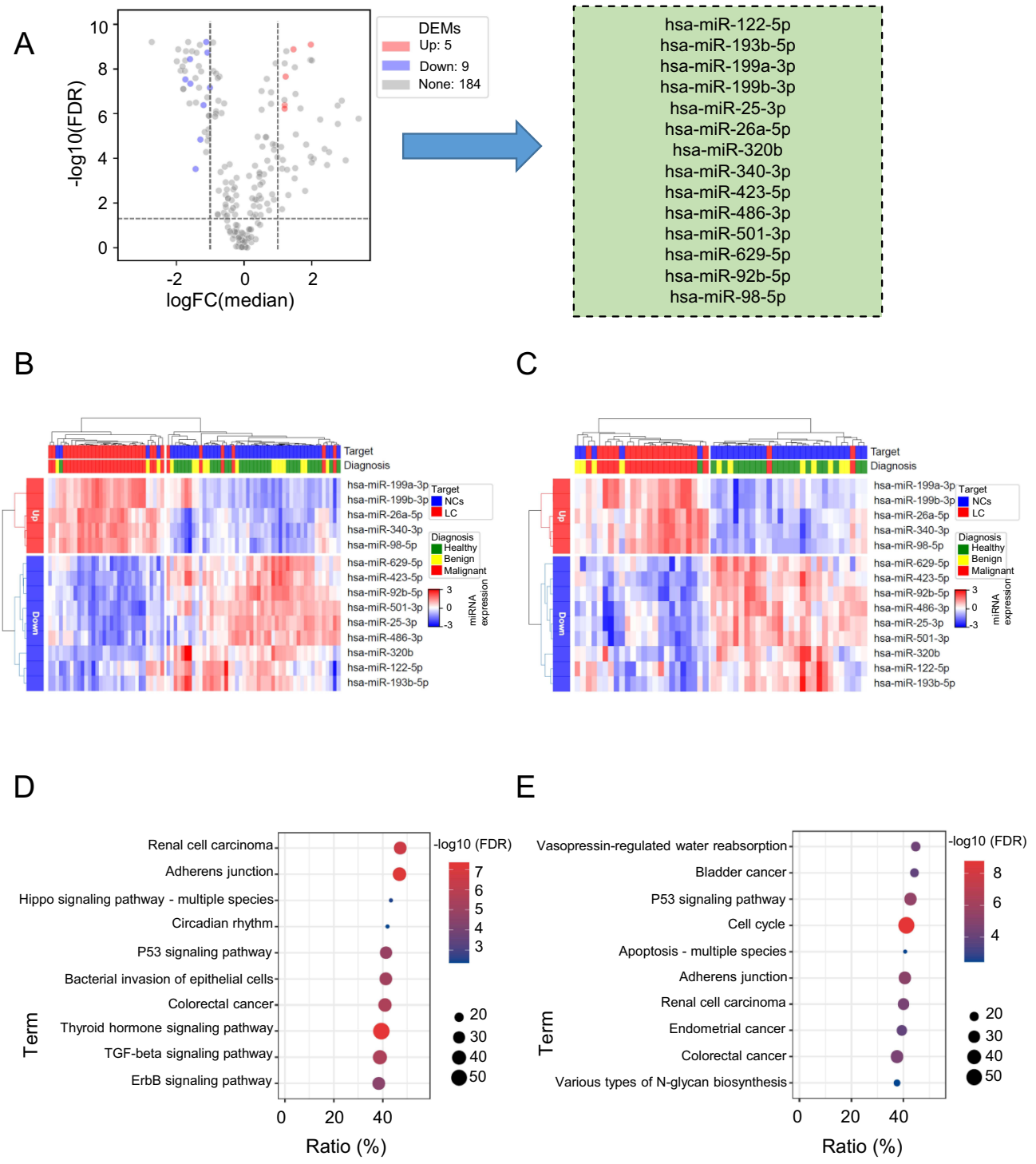


Figure 3 Screening of Candidate sEV miRNA Diagnostic Biomarkers in NCs and Patients with Early-Stage LC. (**A**) Fourteen differentially expressed sEV miRNAs (DEMs) in the NCs and patients with early-stage LC. (**B** and **C**) The clustering results of these 14 candidate DEMs in the training cohort and validation cohort. (**D** and **E**) The results of KEGG pathway enrichment analysis for these 14 DEMs.

found to be effective in distinguishing between NCs and patients with LC in both the training cohort (Figure 3B) and the validation cohort (Figure 3C). By mining the relevant signaling pathways involved with DEMs through KEGG database, we found that up-regulated DEMs act on several cell proliferation regulation and immune-related signaling pathways,

such as the adhesions junction, Hippo, P53, and ERBB signaling pathway, as shown in Figure 3D. Similar biological processes were observed in the down-regulated DEMs, as depicted in Figure 3E.

To identify which sEV miRNAs can be used for early diagnosis of LC, a further screening of diagnostic markers from those 14 DEMs was performed using a circular pipeline with a variety of machine learning algorithms. The feature importance of 14 DEMs was illustrated in Figure 4A. As demonstrated in Figure 4B, the feature importance of the final markers selected by different machine learning models was shown. The diagnostic models were trained based on these markers and evaluated using a 5-fold cross validation (CV) on the training dataset. The receiver operator characteristic (ROC) curves shown in Figure 4C indicated that the LightGBM model (with extra trees) using hsa-miR-423-5p, hsa-miR-340-3p, hsa-miR-320b, hsa-miR-98-5p, hsa-miR-26a-5p, hsa-miR-193b-5p, hsa-miR-629-5p and hsa-miR-92b-5p exhibits the optimal performance with an AUC of 0.956, a sensitivity of 94% and a specificity of 93%.

The generalization performance of these models was assessed on the validation cohort (Figure 4D). The findings indicate that the optimal model also exhibits excellent generalization performance, with an AUC of 0.985, a sensitivity of 86% and a specificity of 97%.

Effects of the Eight Biomarkers on the Signaling Pathway

In order to explore the biological functions of eight biomarkers in the diagnostic model for LC, we mined the relevant signaling pathways in which biomarkers are involved through the KEGG and Reactome databases. The expression levels of the eight biomarkers utilized in the diagnostic model are shown in Figure 5A and Supplementary Figure 1. The three up-regulated biomarkers participate in biological processes such as cell proliferation, differentiation, and apoptosis through various regulatory pathways including adhesions junction, Hippo, P53, and TGF-beta, as demonstrated in Figure 5B. Furthermore, the five down-regulated biomarkers are also associated with cancer-related signaling pathways, as illustrated in Figure 5C. The Reactome database was also utilized to ascertain the potential functional roles of the up-regulated and down-regulated biomarkers, as seen in Figure 5D and E.

Discussion

A substantial body of research has demonstrated the diagnostic efficacy of sEV miRNAs in differentiating tumors. For instance, a scoring model based on seven DEMs achieved a sensitivity of 87.5% and a specificity of 92.3% in differentiating benign from malignant ovarian cancer.[35] A combination of two sEV miRNAs exhibited a sensitivity of 84.1% and a specificity of 96.6% for diagnosing pancreatic ductal adenocarcinoma.[36] Moreover, the diagnostic performance of sEV miRNA has been demonstrated to be effective in distinguishing between liver cancer and non-liver cancer, with a sensitivity of 94.1% and a specificity of 68.4%.[37] The differential diagnostic model developed by Zheng et al showed a specificity of 91.7% in diagnosing indeterminate pulmonary nodules with a diameter of ≤1 cm.[34] The support vector machine model constructed by Zhang et al achieved a sensitivity of 96.9% and a specificity of 100% in diagnosing ground-glass nodules in the lungs.[38] The enormous potential of sEV miRNAs in the early diagnosis of LC is further underlined by our results.

A meta-analysis has revealed that sEV miRNAs have a pooled sensitivity and specificity of 77% and 83%, respectively, for the diagnosis of NSCLC.[39] The diagnostic performance of individual sEV miRNAs is superior to that of serum miRNAs or blood-based tumor markers (CEA, CYFRA21-1, and NSE),[22,40,41] and combinations of sEV miRNAs have been found to outperform individual sEV miRNAs in terms of diagnostic accuracy.[42] Studies with well-performing combinations of sEV miRNAs have reported sensitivities and specificities of 83.8% and 87.1%, respectively, for differentiating early-stage LUAD,[43] and values of 80.25% and 92.31%, respectively, for diagnosing stage I NSCLC patients.[22,44] Existing LC diagnostic models either focus on LUAD or SCC, distinguish between patients and healthy individuals or benign nodules, or lack large-sample validation. We focus on the performance of blood sEV miRNA combinations in complex sample cohorts for early LC diagnosis, with the aim of identifying validated and reliable sEV miRNAs and further exploring their value in LC diagnosis and clinical application.

Our research results indicate that ten sEV miRNAs, namely hsa-miR-423-5p, hsa-miR-340-3p, hsa-miR-320b, hsa-miR-98-5p, hsa-miR-26a-5p, hsa-miR-193b-5p, hsa-miR-629-5p, hsa-miR-92b-5p, hsa-miR-501-3p, and hsa-miR-199b-3p, can be used for early LC through various combinations. Among them, only sEV hsa-miR-320b and hsa-miR-92b-5p have been reported to be used in differentiating SCC and SCLC when combined with other miRNAs.[44,45] Other miRNAs, such as hsa-miR-193b-5p,

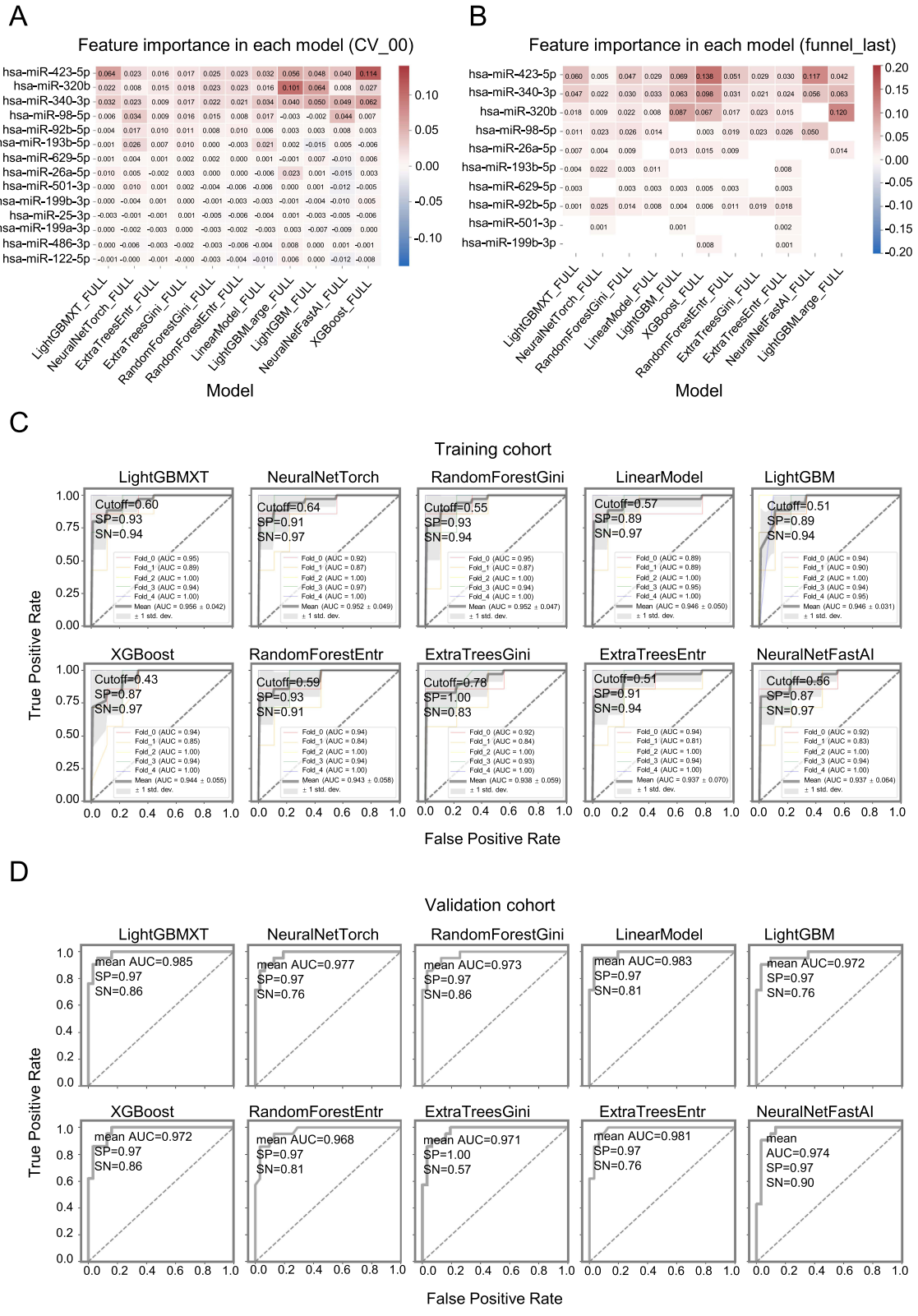


Figure 4 Performance of Plasma sEV miRNA Models in Diagnosing Early-Stage LCs. (**A** and **B**) The permutation feature importance of the 14 candidate DEMs and 10 retained markers across different machine learning models. The color red indicates the marker has a positive effect on model performance, while blue indicates the opposite. (**C** and **D**) ROC curves, sensitivity, and specificity of the diagnostic models for the training and validation cohorts.

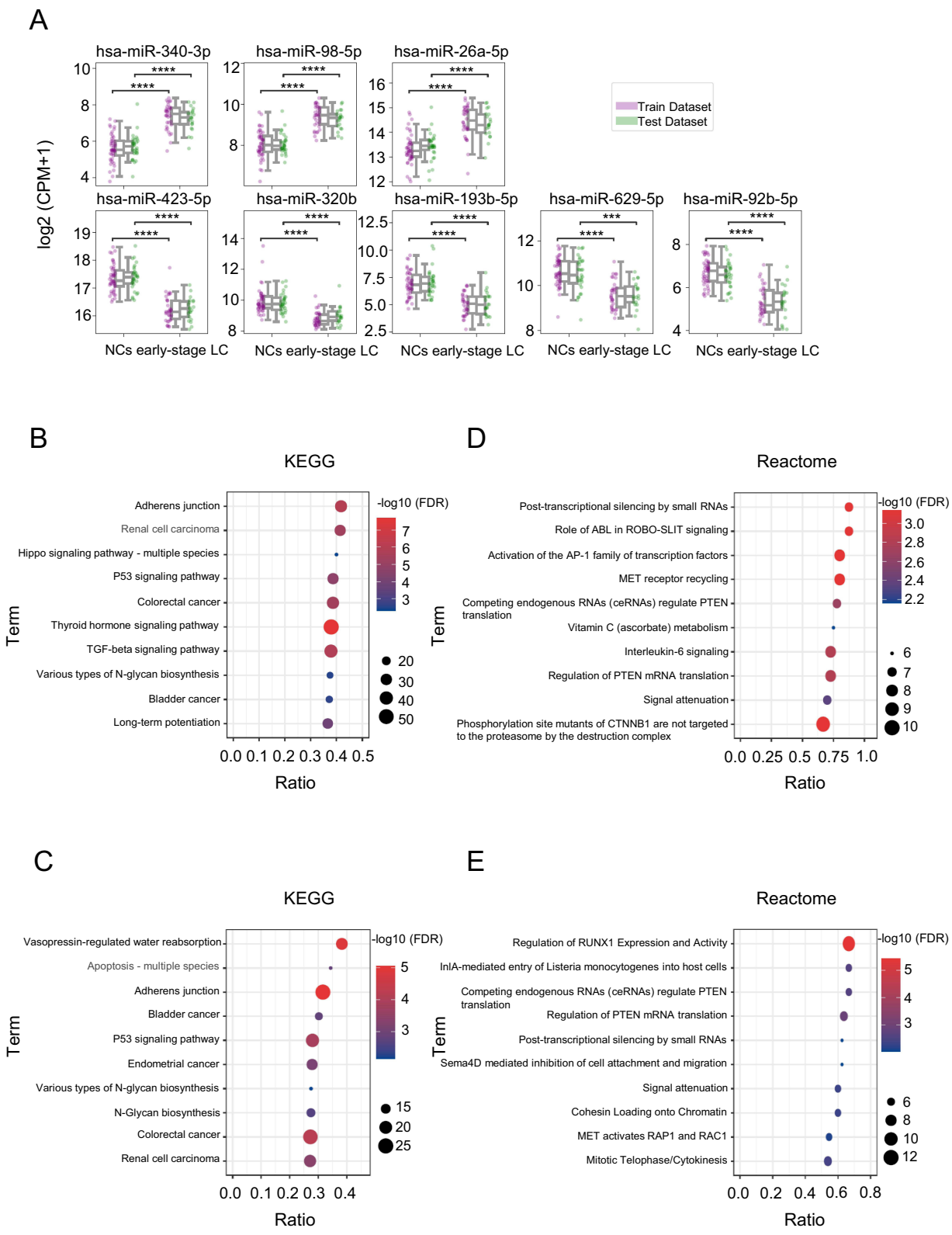


Figure 5 Expression and Biological Functions of sEV miRNA Biomarkers. (**A**) The expression levels of 8 sEV miRNAs used in the diagnostic model between the NCs and patients with early-stage LC. (**B** and **D**) The pathway enrichment analysis results of three up-regulated sEV miRNAs in KEGG and Reactome. (**C** and **E**) The pathway enrichment analysis results of five down-regulated sEV miRNAs in KEGG and Reactome. ***represents that p<0.001, ****represents that p<0.0001.

https://doi.org/10.2147/IJN.S534378

which is reported to be altered only in LUSC, and a panel of five miRNAs including low-expressed hsa-miR-26a-5p in NSCLC, have shown a sensitivity and specificity of 82.5% and 75.1%, respectively, in distinguishing chronic obstructive pulmonary disease (COPD) from NSCLC. Conversely, the up-regulated hsa-miR-423-5p in LC exhibits a relatively poor sensitivity of only 69.1% for LC diagnosis.[46–48] The expression levels of hsa-miR-423-5p, hsa-miR-320b, and hsa-miR-26a-5p, which are important in our diagnostic model, are inconsistent with those reported in previous studies, suggesting that different miRNA signatures exist in populations of different origins and classifications. We need cohort studies that are closer to clinical onset to comprehensively analyze the molecular characteristics of complex populations and facilitate early LC diagnosis.

Our research results indicate that the eight sEV miRNAs involved in the optimal LC diagnostic model participate in biological processes such as cell proliferation, adherens junction, and apoptosis through pathways including P53, TGF-beta, CTNNB1, AP-1, RUNX1, and PTEN, and are implicated in the progression of colorectal cancer, renal cell carcinoma, and bladder cancer. Previous studies have reported that miR-98-5p, miR-26a-5p, miR-629-5p, and miR-320b, can promotes the proliferation and invasive migration of NSCLC through distinct signaling pathways, while miR-340-3p inhibits the proliferation and migration of LUAD cells.[49–53] However, the functionality of miR-423-5p, miR-193b-5p and miR-92b-5p in LC has not been reported. Nevertheless, the down-regulation of miR-423-5p and miR-193b-5p contributes to the enhancement of drug resistance in colorectal cancer cells towards radiotherapy and chemotherapy.[54,55] There have been no reported functional studies concerning adherens junction, thyroid hormone signaling pathway, vasopressin-regulated water reabsorption, and immune regulation. Therefore, we need more research focusing on the mechanistic roles of sEV miRNAs in LC.

It is important to note that the present study is not without its limitations. Firstly, the lack of multi-center samples for validating the expression levels of DEMs in the early LC diagnosis model among healthy individuals, patients with nodules, and lung cancer patients was not employed. This was done in order to eliminate DEMs that were inconsistent with the expression levels obtained by next-generation sequencing. Secondly, the lack of an external validation cohort prevents us from verifying the reproducibility of the early LC diagnosis model constructed in this study. Thirdly, the study is retrospective and lacks prospective large-cohort sample collection to further evaluate the clinical application value of the diagnostic model. Finally, this study only focused on miRNAs, one type of content within sEV, and did not combine other contents or blood-based tumor markers to further enhance the performance of the early LC diagnosis model.

Conclusions

Our finding identified and validated that plasma sEV miRNAs can be used for differentiating between patients with LC from NCs and demonstrated superior diagnostic performance. This suggests that plasma sEV miRNAs are promising and potential liquid biopsy biomarkers for LC detection.

Abbreviations

LC, lung cancer; sEV, small extracellular vesicle; NCs, non-malignant controls; AUC, area under the curve; NSCLC, non-small cell lung cancer; SCLC, small cell lung cancer; LUAD, lung adenocarcinoma; SCC, squamous cell carcinoma; CEA, carcinoembryonic antigen; CYFRA21-1, cytokeratin 19 fragment; NSE, neuron-specific enolase; SCC-Ag, squamous cell carcinoma antigen; proGRP, pro-gastrin-releasing peptide; TEM, transmission electron microscopy; NTA, nanoparticle tracking analysis; TMM, trimmed mean of M-values; KEGG, Kyoto Encyclopedia of Genes and Genomes; DEMs, differentially expressed sEV miRNAs; ML, machine learning; ROC, receiver operator characteristic; CV, cross validation.

Data Sharing Statement

The datasets generated and/or analyzed during the present study are available from the corresponding author (e-mail: maowm@zjcc.org.cn) upon reasonable request.

Ethics Approval Statement

This study was approved by the Ethics Committee of Jiangxi Provincial Cancer Hospital (No. 2021ky231). Our study had adhered to the guidelines and principles stated in the "Declaration of Helsinki".

Patient Consent Statement

Written informed consent was obtained from the NCs and patients with LC for the examination of their samples and the use of their clinical data.

Acknowledgments

We appreciate the support and participation of the physicians and patients in this study. We would like to thank Sheng Chen for her assistance to sample collection and data processing.

Author Contributions

All authors made a significant contribution to the work reported, whether that is in the conception, study design, execution, acquisition of data, analysis and interpretation, or in all these areas; took part in drafting, revising or critically reviewing the article; gave final approval of the version to be published; have agreed on the journal to which the article has been submitted; and agree to be accountable for all aspects of the work.

Funding

This work was supported by the Foundation of Key R&D plan of Jiangxi Province (20212BBG71006).

Disclosure

All of the authors affiliated with 3D Medicines Inc. are current or former employees. No potential competing interests were disclosed by the other authors.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2018;68(6):394–424. doi:10.3322/caac.21492
2. Sung H, Ferlay J, Siegel RL, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2021;71(3):209–249. doi:10.3322/caac.21660
3. Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229–263. doi:10.3322/caac.21834
4. Travis WD, Brambilla E, Burke AP, Marx A, Nicholson AG. Introduction to The 2015 World Health Organization classification of tumors of the lung, pleura, thymus, and heart. *J Thorac Oncol*. 2015;10(9):1240–1242. doi:10.1097/JTO.0000000000000663
5. Walters S, Maringe C, Coleman MP, et al. Lung cancer survival and stage at diagnosis in Australia, Canada, Denmark, Norway, Sweden and the UK: a population-based study, 2004-2007. *Thorax*. 2013;68(6):551–564. doi:10.1136/thoraxjnl-2012-202297
6. Goldstraw P, Chansky K, Crowley J, et al. The IASLC lung cancer staging project: proposals for revision of the TNM stage groupings in the forthcoming (Eighth) Edition of the TNM classification for lung cancer. *J Thorac Oncol*. 2016;11(1):39–51. doi:10.1016/j.jtho.2015.09.009
7. Oncology Society of Chinese Medical Association, Chinese Medical Association Publishing House. [Chinese Medical Association guideline for clinical diagnosis and treatment of lung cancer (2023 edition)]. *Zhonghua Yi Xue Za Zhi*. 2023;103(27):2037–2074. Danish. doi:10.3760/cma.j.cn112137-20230510-00767
8. Nooreldeen R, Bach H. Current and Future Development in Lung Cancer Diagnosis. *Int J Mol Sci*. 2021;22(16):8661. doi:10.3390/ijms22168661
9. Blandin Knight S, Crosbie PA, Balata H, Chudziak J, Hussell T, Dive C. Progress and prospects of early detection in lung cancer. *Open Biol*. 2017;7(9):170070. doi:10.1098/rsob.170070
10. Ren F, Fei Q, Qiu K, Zhang Y, Zhang H, Sun L. Liquid biopsy techniques and lung cancer: diagnosis, monitoring and evaluation. *J Exp Clin Cancer Res*. 2024;43(1):96. doi:10.1186/s13046-024-03026-7
11. Li W, Liu JB, Hou LK, et al. Liquid biopsy in lung cancer: significance in diagnostics, prediction, and treatment monitoring. *Mol Cancer*. 2022;21(1):25. doi:10.1186/s12943-022-01505-z
12. Raez LE, Brice K, Dumais K, et al. Liquid biopsy versus tissue biopsy to determine front line therapy in metastatic non-small cell lung cancer (NSCLC). *Clin Lung Cancer*. 2023;24(2):120–129. doi:10.1016/j.cllc.2022.11.007
13. Korkmaz ET, Koksal D, Aksu F, et al. Triple test with tumor markers CYFRA 21.1, HE4, and ProGRP might contribute to diagnosis and subtyping of lung cancer. *Clin Biochem*. 2018;58:15–19. doi:10.1016/j.clinbiochem.2018.05.001
14. Chen G, Guo P, Zhao H, Zhao D, Yang D. The clinical value of combined detection of seven lung cancer-related autoantibodies in assisting the diagnosis of non-small-cell lung cancer. *Biomarker Med*. 2024;18(20):917–925. doi:10.1080/17520363.2024.2404379
15. Frydrychowicz M, Kuszel Ł, Dworacki G, Budna-Tukan J. MicroRNA in lung cancer-a novel potential way for early diagnosis and therapy. *J Appl Genet*. 2023;64(3):459–477. doi:10.1007/s13353-023-00750-2
16. Zhou X, Wen W, Shan X, et al. A six-microRNA panel in plasma was identified as a potential biomarker for lung adenocarcinoma diagnosis. *Oncotarget*. 2017;8(4):6513–6525. doi:10.18632/oncotarget.14311
17. Liu L, Zhang F, Niu D, Guo X, Lei T, Liu H. Diagnostic value of microRNA-200 expression in peripheral blood-derived extracellular vesicles in early-stage non-small cell lung cancer. *Clin Exp Med*. 2024;24(1):214. doi:10.1007/s10238-024-01455-4

18. Kim M, Park J, Oh S, et al. Deep learning model integrating cfDNA methylation and fragment size profiles for lung cancer diagnosis. *Sci Rep*. 2024;14(1):14797. doi:10.1038/s41598-024-63411-2
19. Tang C, Sun SX, Gu C, et al. Diagnostic and prognostic values of tsRNAs in lung cancer: a meta-analysis. *BMC Cancer*. 2025;25(1):153. doi:10.1186/s12885-025-13536-y
20. Kalluri R, LeBleu VS. The biology, function, and biomedical applications of exosomes. *Science*. 2020;367(6478):eaau6977. doi:10.1126/science.aau6977
21. Yu D, Li Y, Wang M, et al. Exosomes as a new frontier of cancer liquid biopsy. *Mol Cancer*. 2022;21(1):56. doi:10.1186/s12943-022-01509-9
22. He X, Park S, Chen Y, Lee H. Extracellular vesicle-associated miRNAs as a biomarker for lung cancer in liquid biopsy. *Front Mol Biosci*. 2021;8:630718. doi:10.3389/fmolb.2021.630718
23. Liu C, Chen J, Liao J, et al. Plasma extracellular vesicle long RNA in diagnosis and prediction in small cell lung cancer. *Cancers*. 2022;14(22):5493. doi:10.3390/cancers14225493
24. Guo W, Huai Q, Liu T, et al. Plasma extracellular vesicle long RNA profiling identifies a diagnostic signature for stage I lung adenocarcinoma. *Transl Lung Cancer Res*. 2022;11(4):572–587. doi:10.21037/tlcr-21-729
25. Zhang Y, Liu W, Zhang H, et al. Extracellular vesicle long RNA markers of early-stage lung adenocarcinoma. *Int J Cancer*. 2023;152(7):1490–1500. doi:10.1002/ijc.34386
26. Jakobsen KR, Paulsen BS, Bæk R, Varming K, Sorensen BS, Jørgensen MM. Exosomal proteins as potential diagnostic markers in advanced non-small cell lung carcinoma. *J Extracell Vesicles*. 2015;4:26659. doi:10.3402/jev.v4.26659
27. Yuan L, Chen Y, Ke L, et al. Plasma extracellular vesicle phenotyping for the differentiation of early-stage lung cancer and benign lung diseases. *Nanoscale Horiz*. 2023;8(6):746–758. doi:10.1039/d2nh00570k
28. Roman-Canal B, Moiola CP, Gatius S, et al. EV-associated miRNAs from pleural lavage as potential diagnostic biomarkers in lung cancer. *Sci Rep*. 2019;9(1):15057. doi:10.1038/s41598-019-51578-y
29. Martin M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.journal*. 2011;17(1):10. doi:10.14806/ej.17.1.200
30. Harrow J, Frankish A, Gonzalez JM, et al. GENCODE: the reference human genome annotation for The ENCODE Project. *Genome Res*. 2012;22(9):1760–1774. doi:10.1101/gr.135350.111
31. Griffiths-Jones S, Saini HK, van Dongen S, Enright AJ. miRBase: tools for microRNA genomics. *Nucleic Acids Res*. 2008;36(Database issue):D154–158. doi:10.1093/nar/gkm952
32. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res*. 2000;28(1):27–30. doi:10.1093/nar/28.1.27
33. Milacic M, Beavers D, Conley P, et al. The reactome pathway knowledgebase 2024. *Nucleic Acids Res*. 2024;52(D1):D672–D678. doi:10.1093/nar/gkad1025
34. Zheng D, Zhu Y, Zhang J, et al. Identification and evaluation of circulating small extracellular vesicle microRNAs as diagnostic biomarkers for patients with indeterminate pulmonary nodules. *J Nanobiotechnology*. 2022;20(1):172. doi:10.1186/s12951-022-01366-0
35. Li L, Zhang F, Zhang J, et al. Identifying serum small extracellular vesicle MicroRNA as a noninvasive diagnostic and prognostic biomarker for ovarian cancer. *ACS Nano*. 2023;17(19):19197–19210. doi:10.1021/acsnano.3c05694
36. Guo S, Qin H, Liu K, et al. Blood small extracellular vesicles derived miRNAs to differentiate pancreatic ductal adenocarcinoma from chronic pancreatitis. *Clin Transl Med*. 2021;11(9):e520. doi:10.1002/ctm2.520
37. Sheng LQ, Li JR, Qin H, et al. Blood exosomal micro ribonucleic acid profiling reveals the complexity of hepatocellular carcinoma and identifies potential biomarkers for differential diagnosis. *World J Gastrointest Oncol*. 2020;12(10):1195–1208. doi:10.4251/wjgo.v12.i10.1195
38. Zhang JT, Qin H, Man Cheung FK, et al. Plasma extracellular vesicle microRNAs for pulmonary ground-glass nodules. *J Extracell Vesicles*. 2019;8(1):1663666. doi:10.1080/20013078.2019.1663666
39. Huang H, Zhu J, Lin Y, et al. The potential diagnostic value of extracellular vesicle miRNA for human non-small cell lung cancer: a systematic review and meta-analysis. *Expert Rev Mol Diagn*. 2021;21(8):823–836. doi:10.1080/14737159.2021.1935883
40. Wu Q, Yu L, Lin X, et al. Combination of serum miRNAs with serum exosomal miRNAs in early diagnosis for non-small-cell lung cancer. *Cancer Manag Res*. 2020;12:485–495. doi:10.2147/CMAR.S232383
41. Wu Y, Wei J, Zhang W, Xie M, Wang X, Xu J. Serum exosomal miR-1290 is a potential biomarker for lung adenocarcinoma. *Oncol Targets Ther*. 2020;13:7809–7818. doi:10.2147/OTT.S263934
42. Zhong Y, Ding X, Bian Y, et al. Discovery and validation of extracellular vesicle-associated miRNAs as noninvasive detection biomarkers for early-stage non-small-cell lung cancer. *Mol Oncol*. 2021;15(9):2439–2452. doi:10.1002/1878-0261.12889
43. Gao S, Guo W, Liu T, et al. Plasma extracellular vesicle microRNA profiling and the identification of a diagnostic signature for stage I lung adenocarcinoma. *Cancer Sci*. 2022;113(2):648–659. doi:10.1111/cas.15222
44. Jin X, Chen Y, Chen H, et al. Evaluation of tumor-derived exosomal miRNA as potential diagnostic biomarkers for early-stage non-small cell lung cancer using next-generation sequencing. *Clin Cancer Res*. 2017;23(17):5311–5319. doi:10.1158/1078-0432.CCR-17-0577
45. Kim DH, Park H, Choi YJ, et al. Identification of exosomal microRNA panel as diagnostic and prognostic biomarker for small cell lung cancer. *Biomark Res*. 2023;11(1):80. doi:10.1186/s40364-023-00517-1
46. Sun F, Yang X, Jin Y, et al. Bioinformatics analyses of the differences between lung adenocarcinoma and squamous cell carcinoma using The Cancer Genome Atlas expression data. *Mol Med Rep*. 2017;16(1):609–616. doi:10.3892/mmr.2017.6629
47. Leidinger P, Brefort T, Backes C, et al. High-throughput qRT-PCR validation of blood microRNAs in non-small cell lung cancer. *Oncotarget*. 2016;7(4):4611–4623. doi:10.18632/oncotarget.6566
48. Ashirbekov Y, Khamitova N, Satken K, et al. Circulating MicroRNAs as biomarkers for the early diagnosis of lung cancer and its differentiation from tuberculosis. *Diagnostics*. 2024;14(23):2684. doi:10.3390/diagnostics14232684
49. Zhang L, Liang J, Qin H, et al. Lnc AC016727.1/BACH1/HIF-1 α signal loop promotes the progression of non-small cell lung cancer. *J Exp Clin Cancer Res*. 2023;42(1):296. doi:10.1186/s13046-023-02875-y
50. Wang RQ, Long XR, Zhou NN, et al. Lnc-GAN1 expression is associated with good survival and suppresses tumor progression by sponging mir-26a-5p to activate PTEN signaling in non-small cell lung cancer. *J Exp Clin Cancer Res*. 2021;40(1):9. doi:10.1186/s13046-020-01819-0
51. Gong J, Ma L, Peng C, Liu J. LncRNA MAGI2-AS3 acts as a tumor suppressor that attenuates non-small cell lung cancer progression by targeting the miR-629-5p/TXNIP axis. *Ann Transl Med*. 2021;9(24):1793. doi:10.21037/atm-21-6466

52. Zhang S, Zhang X, Sun Q, et al. LncRNA NR2F2-AS1 promotes tumourigenesis through modulating BMI1 expression by targeting miR-320b in non-small cell lung cancer. *J Cell Mol Med*. 2019;23(3):2001–2011. doi:10.1111/jcmm.14102
53. Ren K, Yu Y, Wang X, Liu H, Zhao J. MiR-340-3p-HUS1 axis suppresses proliferation and migration in lung adenocarcinoma cells. *Life Sci*. 2021;274:119330. doi:10.1016/j.lfs.2021.119330
54. Shang Y, Wang L, Zhu Z, et al. Downregulation of miR-423-5p contributes to the radioresistance in colorectal cancer cells. *Front Oncol*. 2020;10:582239. doi:10.3389/fonc.2020.582239
55. Wang J, Zhang X, Zhang J, Chen S, Zhu J, Wang X. Long noncoding RNA CRART16 confers 5-FU resistance in colorectal cancer cells by sponging miR-193b-5p. *Cancer Cell Int*. 2021;21(1):638. doi:10.1186/s12935-021-02353-5

International Journal of Nanomedicine

Dovepress
Taylor & Francis Group

Publish your work in this journal

The International Journal of Nanomedicine is an international, peer-reviewed journal focusing on the application of nanotechnology in diagnostics, therapeutics, and drug delivery systems throughout the biomedical field. This journal is indexed on PubMed Central, MedLine, CAS, SciSearch®, Current Contents®/Clinical Medicine, Journal Citation Reports/Science Edition, EMBase, Scopus and the Elsevier Bibliographic databases. The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit http://www.dovepress.com/testimonials.php to read real quotes from published authors.

Submit your manuscript here: https://www.dovepress.com/international-journal-of-nanomedicine-journal