

Bioinformatics analysis to determine the prognostic value and prospective pathway signaling of miR-126 in non-small cell lung cancer

Zichen Jiao^{1,2#}, Ao Yu^{1,2#}, Xiaofeng He^{1,2#}, Yulong Xuan^{1,2}, He Zhang¹, Guojun Wang³, Minke Shi^{1,2}, Tao Wang^{1,2}

¹Department of Thoracic Surgery, Nanjing Drum Tower Hospital, Medical School of Nanjing University, Nanjing, China; ²Department of Thoracic Surgery, Nanjing Drum Tower Hospital, Nanjing Medical University, Nanjing, China; ³Department of Thoracic Surgery, Jintan People's Hospital, Changzhou, China

Contributions: (I) Conception and design: Z Jiao, A Yu, T Wang; (II) Administrative support: Z Jiao, T Wang; (III) Provision of study materials or patients: Z Jiao, A Yu, T Wang; (IV) Collection and assembly of data: Z Jiao, X He, Y Xuan, H Zhang, G Wang; (V) Data analysis and interpretation: Z Jiao, Y Xuan, H Zhang, G Wang, M Shi; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

[#]These authors contributed equally to this work.

Correspondence to: Tao Wang. Department of Thoracic Surgery, Nanjing Drum Tower Hospital, Medical School of Nanjing University, Nanjing, China. Email: wangtao_pumc@live.cn; Minke Shi. Department of Thoracic Surgery, Nanjing Drum Tower Hospital, Medical School of Nanjing University, Nanjing, China. Email: smk0000001@163.com; Guojun Wang. Department of Thoracic Surgery, Jintan People's Hospital, Changzhou, China. Email: afei3636@sina.com.

Background: MicroRNAs (miRNAs) have been demonstrated to play crucial roles in the initiation and development of non-small cell lung cancer (NSCLC). However, further investigation of the specific role of miR-126 in NSCLC is still required.

Methods: An analysis of miR-126 expression in NSCLC was carried out using the Gene Expression Omnibus (GEO) database, and a literature review was also performed. The differentially expressed genes (DEGs) in three mRNA datasets, GSE18842, GSE19804, and GSE101929, from GEO were identified. Following the prediction of hsa-miR-126-5p target genes by TargetScan, the overlap of miR-126 target genes with DEGs in NSCLC was examined. After that, Gene Ontology enrichment and Kyoto Encyclopedia of Genes and Genomes pathway analyses were performed. Finally, an analysis to identify the impact of hub genes on the prognosis of NSCLC was carried out on the basis of a protein-protein interaction (PPI) network constructed using STRING and Cytoscape.

Results: The data in the literature review revealed a trend that miR126 was downregulated in NSCLC. The number of both NSCLC-related and miR-126-related DEGs was 187. Dozens of DEGs were significantly enriched in biological regulation, cell membrane binding, and signal receptor binding. In the PPI network analysis, 3 of 10 identified hub genes, namely *NCAPG*, *MELK*, and *KIAA0101*, were obviously related to poor prognosis in NSCLC; the survival rate was low among patients with high expression levels of these genes. Furthermore, through network analysis, TPX2, HMMR, and ANLN were identified as recessive miR-126-related genes that may be involved in NSCLC.

Conclusions: MiR-126 plays an essential role in the biological processes of NSCLC through binding to target genes and influences the prognosis of patients with the disease.

Keywords: Non-small cell lung cancer (NSCLC); miR-126; differentially expressed genes (DEGs); functional factors

Submitted Sep 04, 2020. Accepted for publication Dec 16, 2020.

doi: 10.21037/atm-20-7520

View this article at: <http://dx.doi.org/10.21037/atm-20-7520>

Introduction

Lung cancer ranks first among all cancers in terms of global incidence and mortality. Due to its high degree of malignancy, effective treatments for lung cancer are currently lacking (1). In China, the numbers of new cases and deaths associated with lung cancer have gradually increased in recent years (2). The pathogenesis of lung cancer is unclear and its early symptoms vary from patient to patient. As a consequence, a majority of patients are already at the middle or late stages of the disease by the time of histopathological diagnosis, meaning their best opportunity for early surgical intervention is missed. However, even in patients who are able to undergo surgery to achieve complete tumor resection, the rates of recurrence and metastasis during follow-up are often as high as 1/3 (3).

A majority of cases of lung cancer are non-small cell lung cancer (NSCLC). In-depth study and investigation of driver genes in lung cancer have led to an improved understanding of the molecular biology of NSCLC, and therapies targeting driver genes have achieved remarkable results. Targeted therapy, a new therapeutic approach for lung cancer, has achieved an improvement in efficiency and safety, and also has low toxicity, high accuracy, and minimal adverse effects (4,5). In particular, advancements in next-generation sequencing (NGS) technology have led to the promotion of driver gene detection in NSCLC and the subsequent development of targeted therapy (6).

Therefore, the identification of new targets related to the prognosis of NSCLC and further elucidation of the underlying mechanisms would significantly improve the clinical effect of targeted therapy as well as the quality of life for patients with NSCLC.

MicroRNAs (miRNAs), as non-coding small-molecule RNAs, are completely or partially complementary to the three prime untranslated region (3'-UTR), coding sequence (CDS), or promoter region of messenger RNAs (mRNAs). Post-transcriptionally, miRNAs regulate gene expression by inhibiting mRNA translation or directly degrading target mRNA (7). Studies on miRNAs and their various regulatory functions have shown them to play multiple roles in physiological and pathological processes, such as cell differentiation, proliferation, and apoptosis; moreover, their importance in the initiation and development of various tumors has also been reported (8,9). Abnormal expression of miR-126-5p (miR-126) has increasingly been proved to participate in biological processes of multiple cancers. Wei *et al.* reported that miR-126 inhibited the

viability of colorectal cancer cells by hindering apoptosis and autophagy induced by mammalian target of rapamycin (mTOR) (10). In malignant mesothelioma, miRNAs can, through the regulation of epigenetics, serve as biomarkers for early diagnosis and treatment (11). In their study, Kim *et al.* (12) concluded that the production of mouse embryonic stem cell-derived FLK1⁺ cells was regulated by the ETV2/ER71 via miR-126-MAPK pathway. In some studies on the mechanism of miR-126 in the regulation of the malignant biological behavior of NSCLC cells, the results showed that miR-126 could regulate susceptibility to apoptosis, migration, cycle, and proliferation via the STAT3 signaling pathway (13). Moreover, previous studies have confirmed that the expression level of miR-126 is consistently down-regulated in NSCLC and plays a pivotal role in prognosis (14,15).

Therefore, this study aimed to explain the role of miR-126 in NSCLC and to confirm its potential molecular targets through a bioinformatics analysis using Gene Expression Omnibus (GEO) datasets, and a literature review.

We present the following article in accordance with the MDAR reporting checklist (available at <http://dx.doi.org/10.21037/atm-20-7520>).

Methods

The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013).

Selection of GEO datasets

On using GEO database (<http://www.ncbi.nlm.nih.gov/geo/>), the NSCLC gene chip map was available to obtain. The database was searched using the following keywords: (Lung) AND (Neoplasm OR Neoplasia OR Tumor OR Cancer OR Carcinoma OR Malignant OR Malignancy) AND (microRNA OR miRNA OR noncodingRNA OR ncRNA OR small RNA). Three microarray datasets related to the expression of miR-126 in NSCLC tissue samples and adjacent noncancerous lung tissue (ANLT) samples were obtained: GSE101929, GSE18842, and GSE19804.

Literature search for studies on miR-126 in NSCLC

The PubMed, China National Knowledge Infrastructure (CNKI) database, Embase, Web of Science databases were searched from inception to 1/4/2020, with the following keywords: (microRNA OR miRNA OR noncoding RNA

OR ncRNA OR small RNA) and (126 OR 126-5P) AND (Lung) and (Neoplasm OR Neoplasia OR Tumor OR Cancer OR Carcinoma OR Malignancy OR Malignant).

Studies on the expression of miR-126 in NSCLC were included. The exclusion criteria were: case reports, reviews, meta-analyses, nonclinical studies, studies with no control group, and conference summaries.

Gene Ontology enrichment and target gene prediction

There were 32, 46, and 60 NSCLC samples, and 34, 45, and 60 ANLT samples contained in the GSE101929, GSE18842, and GSE19804 datasets, respectively. GPL570 platform: Affymetrix Human Genome U133 plus 2.0 (Affymetrix; Thermo Fisher Science, Inc., Waltham, MA, USA) was adopted to analyze all data.

Differentially expressed genes (DEGs) between the NSCLC and ANLT samples were identified using Limma software package (version 3.6.3) in R/BioManager. By default, the false discovery rate (FDR) method introduced by Benjamini and Hochberg was employed to adjust P values, thus correcting false positive results. The cutoff criteria were: $|\log_2(\text{FC})| > 1$ and $P < 0.05$. The platform annotation file, which was downloaded from the database, served as the basis for converting the probe data in the matrix file into gene symbols.

The target genes of hsa-miR-126-5p were predicted using TargetScan (http://www.targetscan.org/vert_72/). Then, the overlapping miR-126 target genes and DEGs in NSCLC were subjected to Gene Ontology (GO) analysis and were visualized using Binto plugin for Cytoscape (version 3.7.2). In addition, the Database for Annotation, Visualization and Integrated Discovery (DAVID) was employed to perform Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analysis. $\text{FDR} \leq 0.05$ was set as the cutoff criterion. The construction of a protein-protein interaction (PPI) network was carried out with STRING 11.0 (<https://string-db.org/>), and the interactions between genes in the network were visualized using the cytoHubba plugin for Cytoscape (version 3.7.2). The cutoff criterion was a confidence score of $C \geq 0.7$. Finally, molecular complex detection (MCODE) was conducted, and the PPI network modules with degree cutoff = 2, node score cutoff = 0.2, k-core = 2, and maximum depth = 100 were selected.

Survival analysis

The survival rates of 12 types of cancer, including gastric

cancer (n=1,440), ovarian cancer (n=2,190), lung cancer (n=3,452), and breast cancer (n=6,234), were assessed using Kaplan-Meier plotter (www.kmplot.com). The median expression levels of specific genes were determined, and then, based on the median expression, NSCLC patients were grouped into high-expression and low-expression groups. The Kaplan-Meier method was used to analyze the NSCLC patients' survival. In the survival analysis of smokers and non-smokers, specific genes showed statistical significance. Hazard ratio (HR) was used as the effect index of survival risk. The GSE102287 dataset was used to verify whether miR-126 has a correlation with DEGs in NSCLC. This analysis included 62 samples.

Statistical analysis

Data were expressed as mean \pm standard deviation (SD). T-test was adopted for comparisons between two groups. The correlation between the expression level of miR-126 and NSCLC was also analyzed by standardized mean difference (SMD) analysis using Stata 15.0 statistical software. The Mantel-Haenszel formula (fixed-effects model) or the Der Simonian-Laird formula (random-effects model) was used to combine and analyze different GEO datasets. With $P \leq 0.05$ or $I^2 \geq 50\%$ representing a significant Q statistic, a random-effects model was adopted; otherwise, a fixed-effects model was used. Spearman's rank correlation was used to analyze the correlation between the expression of DEGs and the levels of miR-126. $P < 0.05$ was considered statistically significant.

Results

miR-126 expression in NSCLC based on GEO

Based on four GEO datasets (GSE102286, GSE29248, GSE63805, and GSE27705), the expression of miR-126 in NSCLC tissue and ANLT was evaluated (*Figure 1*). Compared with that in ANLT samples, miR-126 expression was significantly up-regulated in NSCLC tissues in the GSE102286, GSE29248, and GSE27705 datasets ($P < 0.01$), but the difference was not significant in GSE63805. *Table 1* and *Figure 2A,B,C,D* show the characteristics of the GEO datasets. In accordance with all the included GEO datasets, a statistically significant difference was identified in miR-126 expression between NSCLC and normal tissues (SMD = -1.69, 95% CI: -2.97 to -0.41, $P < 0.01$). A forest map is shown in *Figure 3*.

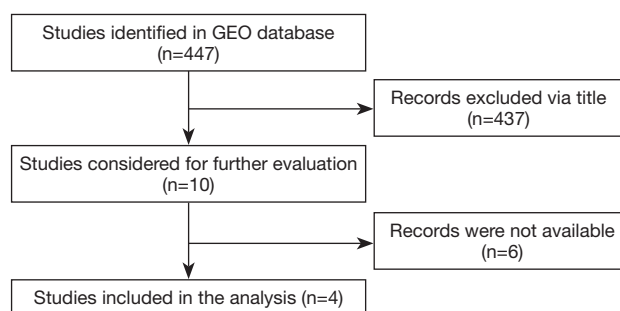


Figure 1 Flow-chart of the research and screening processes for the Gene Expression Omnibus datasets.

Table 1 Characteristics of the Gene Expression Omnibus (GEO) datasets

Dataset	NSCLC			Normal control			t	P
	Mean	SD	n	Mean	SD	n		
GSE27705	0.45	0.22	20	2.07	1.27	10	4.00	0.003
GSE63805	13.79	1.19	31	14.15	1.09	31	1.24	0.220
GSE29248	13.18	0.23	6	13.58	0.19	6	3.31	0.008
GSE102286	12.59	0.72	90	14.34	0.71	89	16.29	0.000
Total	SMD = -1.69 (95% CI: -2.97 to -0.41), P=0.009; I ² =93.3%, P=0.000							

NSCLC, non-small cell lung cancer; SD, standard deviation; SMD, standardized mean difference.

Expression profiles of miR-126 in NSCLC and ANLT in the literature

A literature review of studies on the expression of miR-126 in NSCLC was performed. As shown in *Figure 4*, six studies from the literature that met the selection criteria were selected (14-19). As shown in *Table 2*, miR-126 expression was significantly decreased in NSCLC tissues in comparison with normal tissues.

miR-126 target prediction and bioinformatics analysis, data preprocessing and screening of DEGs

In the GSE18842, GSE19804, and GSE101929 datasets, 1829, 804, and 1921 DEGs were identified, respectively; 595 common DEGs were obtained from these three datasets using Venny 2.1.0 (<https://bioinfogp.cnb.csic.es/tools/venny/>) (*Figures 5,6*). Then, based on TargetScan, 5591 TG_hsa-miR-126-5p were reported, 187 of which overlapped with 595 common DEGs. As could be seen from the GSE101929 data, among the genes associated with hsa-miR-126-5p in NSCLC tissues, 46 were up-regulated and 141 were down-regulated (*Figure 6*). Subsequently, the top 10 up-regulated and down-regulated genes were obtained and are shown in *Table 3*.

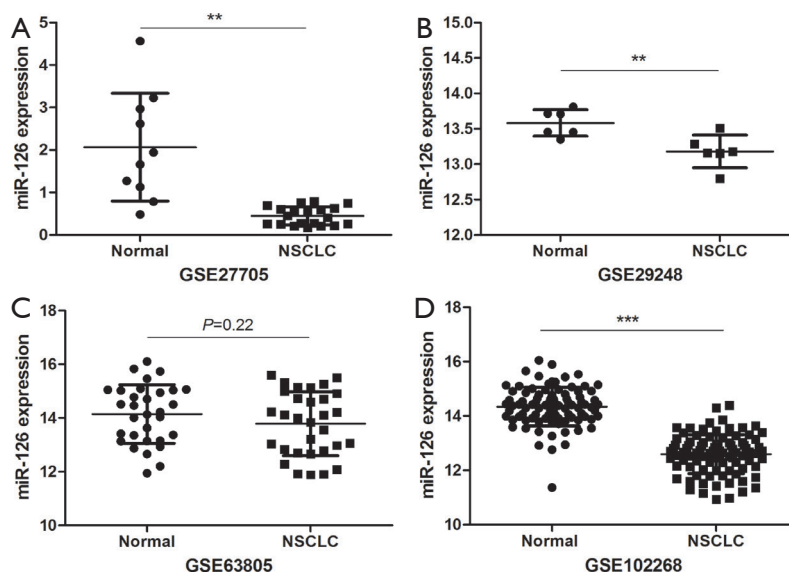


Figure 2 Expression of miR-126 in NSCLC and normal lung tissues in the GEO datasets. (A) miR-126 expression and GSE27705; (B) miR-126 expression and GSE 29248; (C) miR-126 expression and GSE63805; (D) miR-126 and GSE102268. **, P<0.01; ***, P<0.001. NSCLC, non-small cell lung cancer; GEO, Gene Expression Omnibus; Normal, normal lung tissue; miR-126, hsa-miR-126-5p.

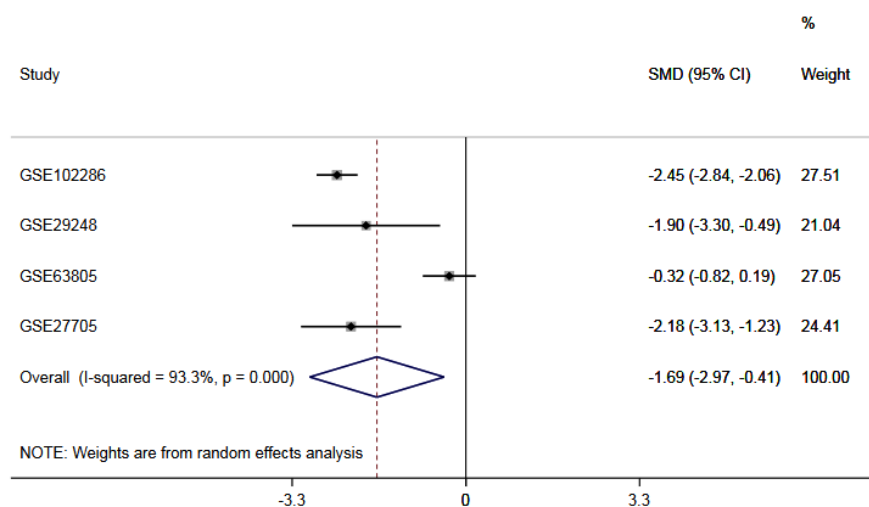


Figure 3 Forest plot of differential expression of miR-126-5p in the GEO datasets. GEO, Gene Expression Omnibus.

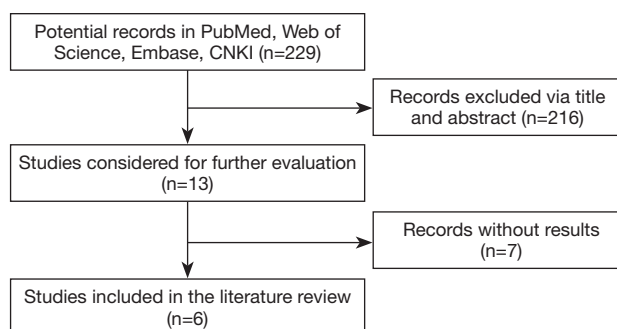


Figure 4 Screening of literature on the expression of miR-126-5p in NSCLC. NSCLC, non-small cell lung cancer.

Functional analysis of miR-126-related DEGs in NSCLC

GO enrichment analysis was performed using Bingo plugin in Cytoscape, and KEGG enrichment analysis was carried out using DAVID. The results of GO enrichment analysis are shown in *Figure 7* and *Table 4*. It was found that many target genes participate in biological processes, mainly including biological regulation, signal receptor binding, cell membrane binding, stimulus response, and glycosaminoglycan binding (*Figure 7*, *Table 4*). In addition, KEGG pathway enrichment analysis (*Table 4*) revealed enrichment in the protein digestion and absorption

Table 2 Studies included from the literature

Author	Year	Country	Case		Control		Result	Detection methods
			Source name	n	Source name	n		
Yang JY	2012	China	NSCLC	167	ANLT	167	Down-regulation	qRT-PCR
Kim	2014	Korea	NSCLC	72	BLT	30	Down-regulation	qRT-PCR
Li XG	2014	China	NSCLC	49	ANLT	49	Down-regulation	qRT-PCR
Chen SW	2015	China	NSCLC	113	ANLT	113	Down-regulation	qRT-PCR
Begum	2015	USA	NSCLC	18	ANLT	18	Down-regulation	qRT-PCR
Switlik	2017	Poland	NSCLC	50	ANLT	50	Down-regulation	qRT-PCR

ANLT, adjacent noncancerous lung tissues; BLT, benign lung tissue; NSCLC, non-small cell lung cancer; qRT-PCR, quantitative real-time polymerase chain reaction.

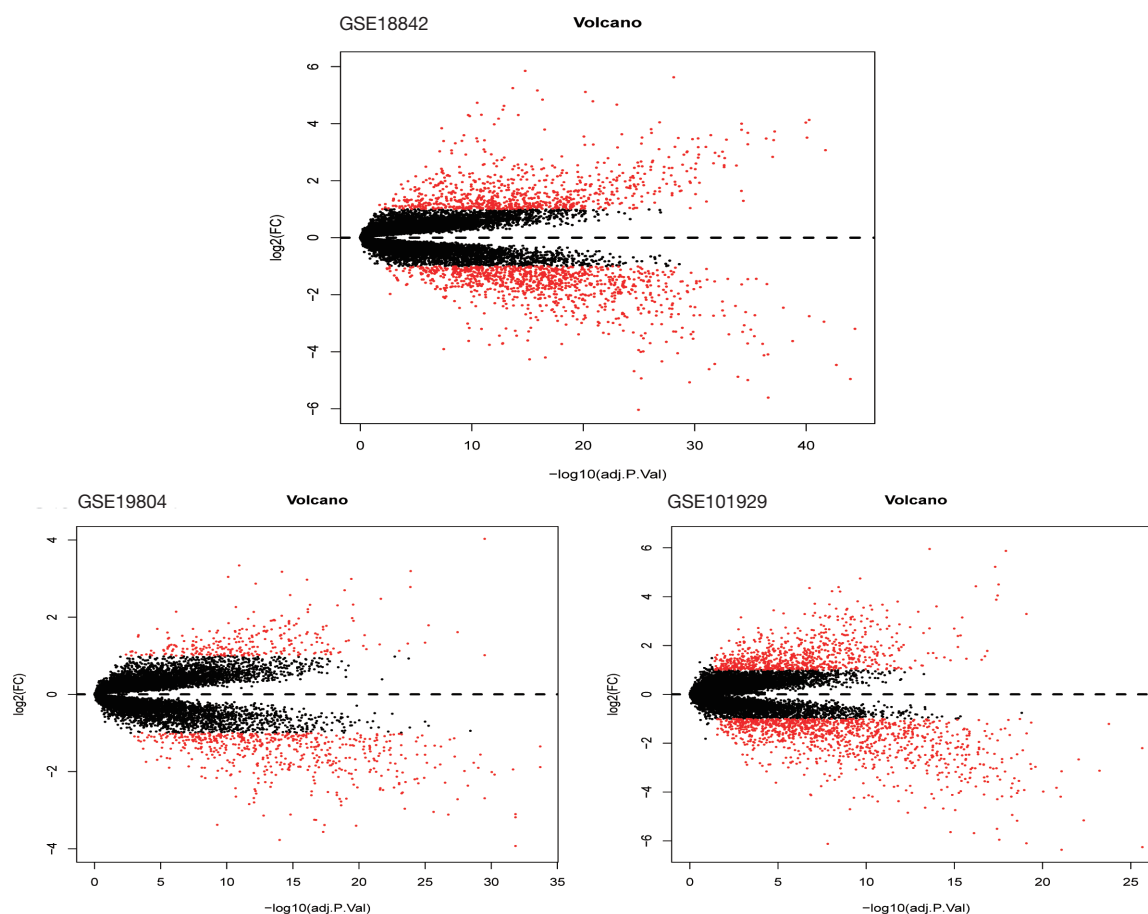


Figure 5 Volcano plot of genome-wide mRNA detected from three NSCLC-related datasets from GEO. Red plots: aberrantly expressed mRNAs with $P < 0.05$ and $|\log_2(FC)| > 1$; red plots above the line: up-regulated genes; red plots below the line: down-regulated genes. Black plots: normally expressed mRNAs. The x-axis represents the adj.P of $-\log_{10}$. P value of each mRNA represents the associated strength. NSCLC, non-small cell lung cancer; GEO, Gene Expression Omnibus; adj.PValue, adjusted P value; FC, fold change.

interaction pathway.

PPI network construction and module selection

A PPI network of miR-126-related DEGs was constructed, with 187 nodes and 281 edges (Figure 8). As shown in Figure 9A, with the top 10 hub genes ranked using degree as the criterion, a close association among hub genes was identified. Following that, MCODE was adopted to obtain an important module (Figure 9B) from the above PPI network, with 12 nodes and 64 edges.

Survival analysis

CCNB1, *KIAA0101*, *BUB1B*, *TPX2*, *NUF2*, *NCAPG*,

MELK, *KIF15*, *HMMR*, and *ANLN* were the top 10 hub genes in the PPI network. Their prognostic values were assessed using Kaplan-Meier plotter. Analysis of the overall survival of NSCLC patients was performed based on high and low expression of each hub gene. The results of overall survival analysis showed that all 10 genes were related to the prognosis of NSCLC. However, the results were inconsistent when NSCLC patients were divided into smokers and non-smokers. Only three genes showed statistical differences between the smokers and non-smokers, namely *NCAPG* (HR = 1.59, 95% CI: 1.40–1.80, $P < 0.01$), *MELK* (HR = 1.6, 95% CI: 1.41–1.82, $P < 0.01$), and *KIAA0101* (HR = 1.56, 95% CI: 1.37–1.77, $P < 0.01$), indicating the association of these three genes with a decreased overall survival rate in NSCLC patients

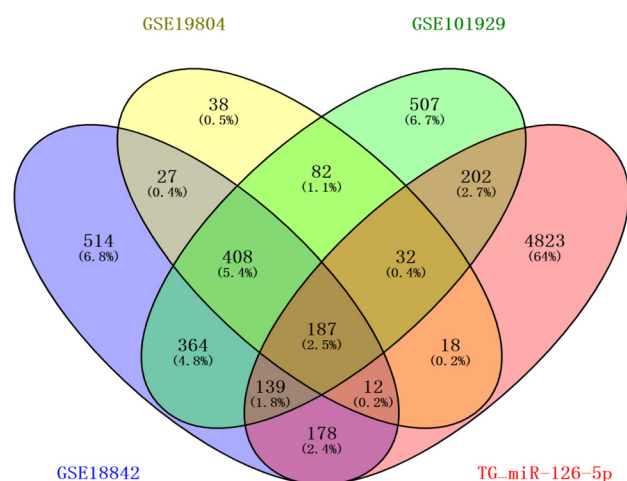


Figure 6 Venn diagram of differentially expressed genes related to hsa-miR-126-5p in TG_miR-126-5p and other three GEO datasets; the overlapping region represents the recognized DEGs. DEGs, differentially expressed genes; TG_miR-126-5p, target genes of hsa-miRNA-126-5p.

(Figure 10). As can be seen in Table 5, the data obtained from the PPI network analysis revealed a negative correlation between miR-126 and the top 10 hub genes, especially ANLN, HMMR, TPX2, and HMMR; these genes might play a role in NSCLC by interacting with miR-126.

Discussion

Based on the datasets from GEO, data from the literature review, and comparison of miRNA expression profiles in NSCLC and normal tissues, the current study revealed a correlation of abnormal miR-126 expression and NSCLC. Moreover, GO, KEGG, PPI network, and Kaplan-Meier survival analyses were employed to identify and analyze the novel markers and potential targets of miR-126, which were involved in the key biological processes in NSCLC.

To date, few studies have investigated miR-126 characteristics in NSCLC; however, down-regulation of miR-126 in NSCLC has been confirmed in various clinical studies. For instance, as reported by Wang *et al.* (20) in 2015, early-stage NSCLC patients had lower miR-126 expression in the blood, which was consistent with findings of Zhu *et al.* (21) and Shang *et al.* (22). Therefore, miR-126 is considered to be a diagnostic marker of early-stage NSCLC.

In NSCLC patients, the level of miR-126 is not only decreased in body fluids, but it is also lower in tumor tissues

Table 3 Top 10 differentially expressed genes related to hsa-miR-126-5p in NSCLC based on data from Gene Expression Omnibus (GSE101929)

DEG	logFC	P value	adj. P value
Up-regulated genes			
<i>COL11A1</i>	4.422953836	2.39E-19	6.08E-17
<i>TMPRSS4</i>	4.221629522	4.37E-10	8.69E-09
<i>TPX2</i>	3.807505294	9.78E-12	3.07E-10
<i>GJB2</i>	3.796664393	2.78E-14	1.66E-12
<i>ADAMDEC1</i>	3.708366029	5.52E-09	8.22E-08
<i>GCNT3</i>	3.685815147	6.40E-10	1.22E-08
<i>MELK</i>	3.59921375	9.24E-17	1.04E-14
<i>ANLN</i>	3.459272472	5.15E-11	1.31E-09
<i>GINS1</i>	3.379747077	1.33E-15	1.12E-13
<i>LGSN</i>	3.311053254	5.50E-07	4.85E-06
Down-regulated genes			
<i>SCN4B</i>	-3.966201158	2.47E-18	4.37E-16
<i>PGC</i>	-4.101979531	1.07E-08	1.46E-07
<i>TCF21</i>	-4.1755375	3.80E-21	1.98E-18
<i>CYP4B1</i>	-4.188821121	1.11E-12	4.46E-11
<i>IGSF10</i>	-4.295135221	8.70E-21	3.70E-18
<i>MAMDC2</i>	-4.641650221	4.93E-16	4.54E-14
<i>GKN2</i>	-4.84796114	6.20E-15	4.38E-13
<i>SOSTDC1</i>	-5.507636397	9.25E-21	3.85E-18
<i>TMEM100</i>	-5.954176691	6.04E-21	2.86E-18
<i>SLC6A4</i>	-6.365475914	2.88E-25	8.40E-22

DEG, differentially expressed gene; FC, fold-change; NSCLC, non-small cell lung cancer.

than in the adjacent normal tissues (14,15); a significant association has been identified between decreased miR-126 expression and NSCLC prognosis (16,17). A significant decrease of miR-126 expression was found in NSCLC tissues in three out of the four GSE datasets used in this study; a statistically significant combined SMD was obtained from the meta-analysis. Therefore, it is of great significance to further study the mechanism of miR-126, so as to identify its target genes and potential therapeutic targets for NSCLC.

MiR-126, as a crucial mRNA, participates in NSCLC regulation. A meta-analysis by Zheng *et al.* (23), for

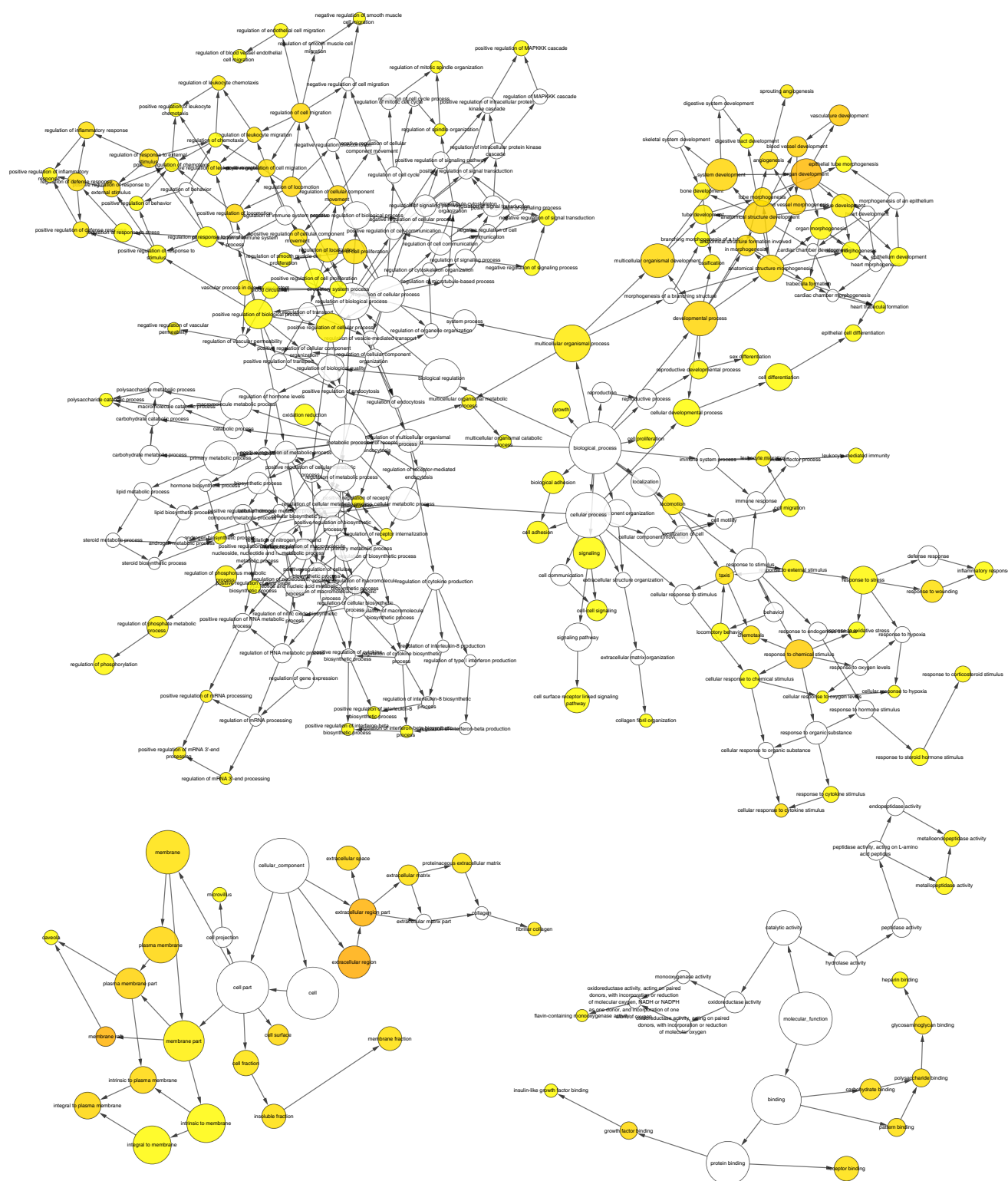


Figure 7 Results of GO enrichment analysis of miR-126 target genes in NSCLC. The yellow circle represents functional enrichment; a larger circle and a darker color show that more genes are enriched in this pathway. The connecting lines represent the associations between genes. NSCLC, non-small cell lung cancer.

Table 4 GO and KEGG enrichment analysis of differentially expressed genes related to hsa-miR-126-5p in NSCLC

Term	Description	Count	FDR
Biological processes			
GO:0050896	Response to stimulus	101	0.0043
GO:0065007	Biological regulation	133	0.0276
GO:0009987	Cellular process	157	0.0408
Molecular function (GO)			
GO:1901681	Sulfur compound binding	10	0.0221
GO:0005539	Glycosaminoglycan binding	11	0.0073
GO:0005102	Signaling receptor binding	30	0.0221
Cellular component			
GO:0005886	Plasma membrane	86	1.58E-06
GO:0071944	Cell periphery	88	1.38E-06
GO:0016020	Membrane	107	0.0024
KEGG pathway			
hsa04974	Protein digestion and absorption	7	0.0061

NSCLC, non-small cell lung cancer; FDR, false discovery rate; KEGG, Kyoto Encyclopedia of Genes and Genomes.

instance, pointed out a positive relationship of increased NSCLC overall survival and high expression of miR-126. MiR-126-3p can hamper the growth, migration, and invasion of NSCLC cells by targeting the CCR1 gene (24). The enzyme activity of mitochondrial respiration and Malate Dehydrogenase 1 (MDH1) in NSCLC cells can be inhibited by miR-126-5p, thereby resulting in cell death (25). NSCLC cell growth is promoted by up-regulated long-chain non-coding RNA (lncRNA) MINCR through a negative regulatory relationship with miR-126/SLC7A5 (26). Guo *et al.* (27) also found a regulatory effect of lncRNA PRNCR1 on NSCLC cell invasion, migration, apoptosis, and proliferation via the PRNCR1/miR-126-5p/MTDH axis. In view of the current situation, further investigation of the molecular mechanisms and clinical value of abnormal miR-126 expression in NSCLC is required. The high correlation of 10 hub genes was revealed in PPI network analysis, and the association of these genes and NSCLC prognosis was obtained from the survival analysis. However, another analysis of NSCLC prognosis was carried

out in smokers and non-smokers due to the significant effect of smoking on lung cancer. After stratified analysis, some genes were no longer significant, and only the expression levels of three genes (*NCAPG*, *MELK*, and *KIAA0101*) were significantly related to the prognosis of NSCLC, regardless of smoking status; high expression of these three genes was associated with a poor NSCLC prognosis. This result also confirms the negative regulatory effect of miR-126.

NCAPG (non-SMC condensin I complex, subunit G), a subunit of the agglutinin complex, participates in the condensation and stabilization of chromosomes in the processes of mitosis and meiosis (28). In renal cell carcinoma, *NCAPG* can be used as a candidate biomarker (29). *MELK* (maternal embryo leucine zipper kinase), a member of the CAMK serine/threonine protein kinase superfamily, is most active in the process of mitosis. Zang *et al.* reported that a high expression of *MELK* leads to poor prognosis in NSCLC (30); our study has reached the same conclusion. *KIAA0101* mainly plays a role in the liver, but its relationship with other organs has been studied less. In the study of Kato *et al.* (31), overexpression of *KIAA0101* was shown to be a predictor of poor prognosis in primary lung cancer, which is similar to our results. Overall, these three genes have the possibility to be novel prognostic markers for NSCLC.

The new candidate target genes of miR-126 in our study, such as *TPX2*, *HMMR*, and *ANLN*, were found to be involved in the regulation of key biological processes of NSCLC. *TPX2* was reported to have a relationship with poor NSCLC prognosis in the study of Schneider *et al.* (32). *HMMR* is identified as a polyhedral molecule, which has a tumor-promoting effect. The product of *HMMR*, hyaluronic acid-mediated cell migration receptor (*RHAMM*), also promotes the occurrence and development of tumors. Specifically, *RHAMM* promotes the migration of tumor cells and has a correlation with tumor pathological stage (33). The expression of A549 anillin (*ANLN*) gene in human lung adenocarcinoma cells is suppressed by carbon ion irradiation via activation of the P13K/Akt signaling pathway. *ANLN* has been demonstrated to be inhibited through the P13K/Akt signaling pathway (34). Negative correlations with significant differences were revealed between miR-126 and *TPX2*, *HMMR*, and *ANLN* in NSCLC in our study, suggesting their important biological regulatory role in NSCLC.

In summary, miR-126 has a critical biological role in NSCLC. However, the pathogenesis and the effect of the

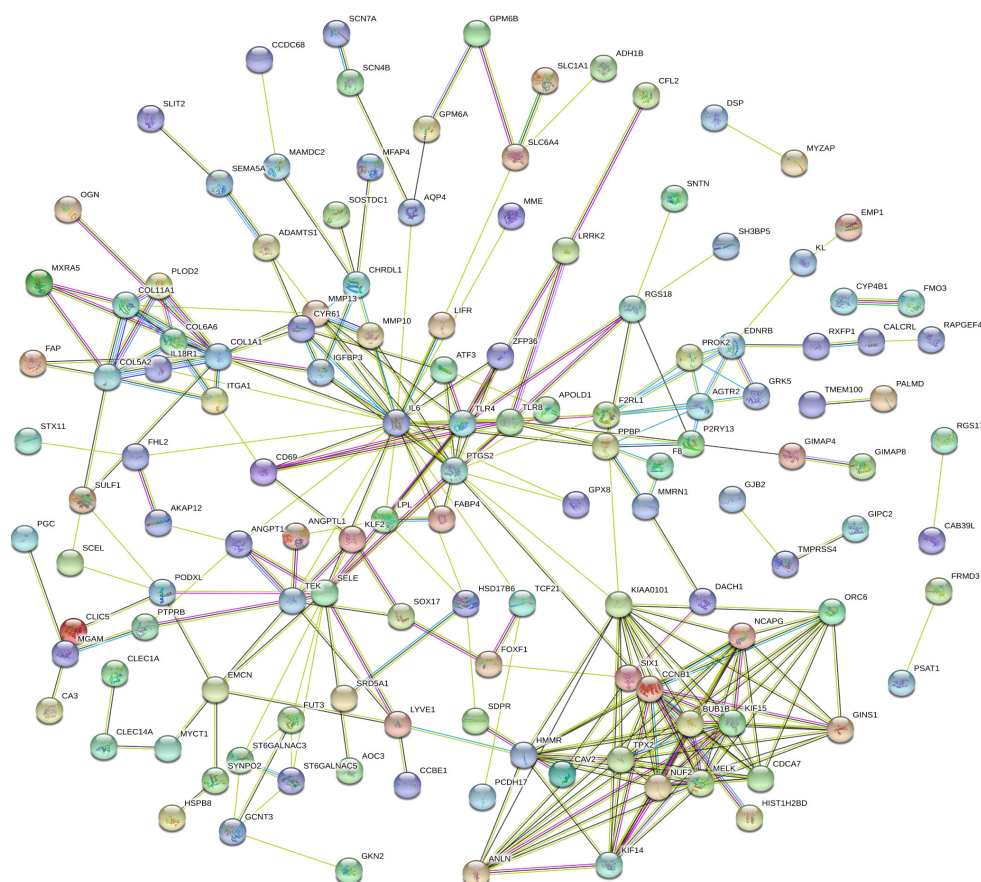


Figure 8 Protein-protein interaction network of hsa-miR-126-5p-related differentially expressed genes. The lines represent interaction between nodes, that is, interaction between proteins. Colored nodes: query proteins and first shell of interactors; white nodes: second shell of interactors; empty nodes: proteins of unknown 3D structure; filled nodes: some 3D structures known or predicted. DEGs, differentially expressed genes.

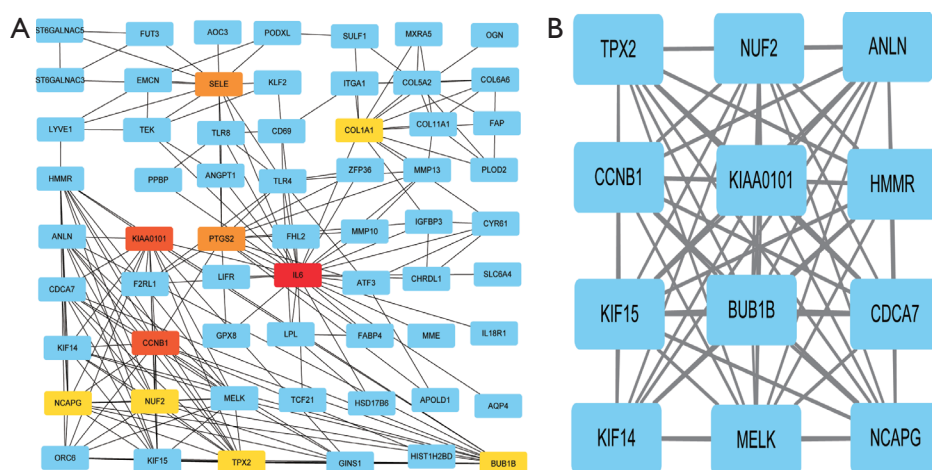


Figure 9 Protein-protein interaction network. (A) Protein-protein interaction network of hub genes of hsa-miR-126-5p-related DEGs; (B) a significant module selected from the protein-protein interaction network of hsa-miR-126-5p-related DEGs. Red and yellow represent the top 10 hub genes. The darker the color, the stronger the association with other genes in the PPI network. The lines represent interaction relationship between nodes. DEGs, differentially expressed genes.

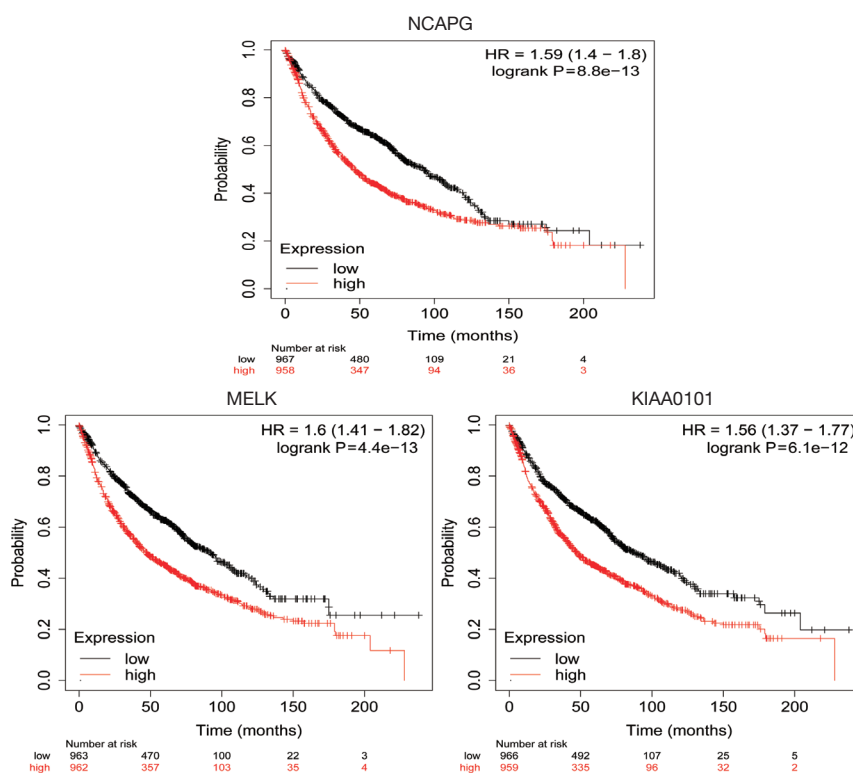


Figure 10 Expression of NCAPG, MELK, KIAA0101 and NSCLC prognosis. According to the median expression level, NSCLC patients are grouped according to low expression and high expression. NSCLC, non-small cell lung cancer.

Table 5 Correlation between hsa-miR-126 and target genes in NSCLC patients (n=62)

miRNA	Target gene	r	P value
hsa-miR-126	<i>ANLN</i>	-0.2886	0.0229
hsa-miR-126	<i>HMMR</i>	-0.2737	0.0314
hsa-miR-126	<i>KIF15</i>	-0.1640	0.2029
hsa-miR-126	<i>MELK</i>	-0.2202	0.0855
hsa-miR-126	<i>NCAPG</i>	-0.2256	0.0780
hsa-miR-126	<i>NUF2</i>	-0.2328	0.0686
hsa-miR-126	<i>TPX2</i>	-0.2602	0.0411
hsa-miR-126	<i>BUB1B</i>	-0.2408	0.0594
hsa-miR-126	<i>KIAA0101</i>	-0.2454	0.0545
hsa-miR-126	<i>CCNB1</i>	-0.2158	0.0921

NSCLC, non-small cell lung cancer.

molecular network regulated by miR-126 in this disease still needs to be determined by further *in vitro* and *in vivo* experiments in the future.

Acknowledgments

Funding: Our work is supported by Natural Science Foundation of Jiangsu Province, Grant/Award No. BK20140601.

Footnote

Reporting Checklist: The authors have completed the MDAR reporting checklist. Available at <http://dx.doi.org/10.21037/atm-20-7520>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <http://dx.doi.org/10.21037/atm-20-7520>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as

revised in 2013).

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2018. *Ca A Cancer Journal for Clinicians* 2018;68:7-30.
2. Parascandola M, Xiao L. Tobacco and the lung cancer epidemic in China. *Transl Lung Cancer Res* 2019;8:S21-30.
3. Ahn S, Jeong JY, Kim HW, et al. Robotic lobectomy for lung cancer: initial experience of a single institution in Korea. *Ann Cardiothorac Surg* 2019;8:226-32.
4. Sacco PC, Maione P, Guida C, et al. The combination of new immunotherapy and radiotherapy: a new potential treatment for locally advanced non-small cell lung cancer. *Curr Clin Pharmacol* 2017;12:4-10.
5. Bagley SJ, Talento S, Mitra N, et al. Comparative effectiveness of carboplatin/pemetrexed with versus without bevacizumab for advanced nonsquamous non-small cell lung cancer. *J Natl Compr Canc Netw* 2019;17:469-77.
6. Helman E, Nguyen M, Karlovich CA, et al. Cell-Free DNA Next-Generation Sequencing Prediction of Response and Resistance to Third-Generation EGFR Inhibitor. *Clin Lung Cancer* 2018;19:518-530.e7.
7. Ma N, Wei T, Wang B, et al. MicroRNA-142-3p inhibits IFN-gamma production via targeting of RICTOR in *Aspergillus fumigatus* activated CD4(+) T cells. *Ann Transl Med* 2019;7:649.
8. Fan J, Xu W, Nan S, et al. MicroRNA-384-5p Promotes Endothelial Progenitor Cell Proliferation and Angiogenesis in Cerebral Ischemic Stroke through the Delta-Like Ligand 4-Mediated Notch Signaling Pathway. *Cerebrovasc Dis* 2020;49:39-54.
9. Jacques C, Tesfaye R, Lavaud M, et al. Implication of the p53-Related miR-34c, -125b, and -203 in the Osteoblastic Differentiation and the Malignant Transformation of Bone Sarcomas. *Cells* 2020;9:810.
10. Wei L, Chen Z, Cheng N, et al. MicroRNA-126 Inhibit Viability of Colorectal Cancer Cell by Repressing mTOR Induced Apoptosis and Autophagy. *Onco Targets Ther* 2020;13:2459-68.
11. Tomasetti M, Gaetani S, Monaco F, et al. Epigenetic Regulation of miRNA Expression in Malignant Mesothelioma: miRNAs as Biomarkers of Early Diagnosis and Therapy. *Front Oncol* 2019;9:1293.
12. Kim JY, Lee DH, Kim JK, et al. ETV2/ER71 regulates the generation of FLK1(+) cells from mouse embryonic stem cells through miR-126-MAPK signaling. *Stem Cell Res Ther* 2019;10:328.
13. Zhang Z, Wang J, Cheng J, et al. Effects of miR-126 on the STAT3 signaling pathway and the regulation of malignant behavior in lung cancer cells. *Oncol Lett* 2018;15:8412-6.
14. Begum S, Hayashi M, Ogawa T, et al. An integrated genome-wide approach to discover deregulated microRNAs in non-small cell lung cancer: Clinical significance of miR-23b-3p deregulation. *Sci Rep* 2015;5:13236.
15. Świtlik W, Karbownik MS, Suwalski M, et al. miR-30a-5p together with miR-210-3p as a promising biomarker for non-small cell lung cancer: A preliminary study. *Cancer Biomark* 2018;21:479-88.
16. Yang J, Lan H, Huang X, et al. MicroRNA-126 Inhibits Tumor Cell Growth and Its Expression Level Correlates with Poor Survival in Non-Small Cell Lung Cancer Patients. *PLoS One* 2012;7:e42978.
17. Kim MK, Jung SB, Kim JS, et al. Expression of microRNA miR-126 and miR-200c is associated with prognosis in patients with non-small cell lung cancer. *Virchows Arch* 2014;465:463-71.
18. Li XG, Wan GS, Liang YX, et al. Effect of miR-126 on cell cycle regulation and prognosis of lung cancer patients. *Journal of Practical Oncology* 2014;29:440-5.
19. Chen SW, Wang TB, Tian YH, et al. Down-regulation of microRNA-126 and microRNA-133b acts as novel predictor biomarkers in progression and metastasis of non small cell lung cancer. *Int J Clin Exp Pathol* 2015;8:14983-8.
20. Wang P, Yang D, Zhang H, et al. Early Detection of Lung Cancer in Serum by a Panel of MicroRNA Biomarkers. *Clin Lung Cancer* 2015;16:313-9.e1.
21. Zhu W, Zhou K, Zha Y, et al. Diagnostic Value of Serum miR-182, miR-183, miR-210, and miR-126 Levels in Patients with Early-Stage Non-Small Cell Lung Cancer. *PLoS One* 2016;11:e0153046.
22. Shang AQ, Xie YN, Wang J, et al. Predicative values of

- serum microRNA-22 and microRNA-126 levels for non-small cell lung cancer development and metastasis: a case-control study. *Neoplasma* 2017;64:453-9.
23. Zheng W, Zhou Y, Lu J, et al. The prognostic value of miR-126 expression in non-small-cell lung cancer: a meta-analysis. *Cancer Cell Int* 2017;17:71.
 24. Liu R, Zhang YS, Zhang S, et al. MiR-126-3p suppresses the growth, migration and invasion of NSCLC via targeting CCR1. *Eur Rev Med Pharmacol Sci* 2019;23:679-89.
 25. Lima Queiroz A, Zhang B, Comstock DE, et al. miR-126-5p targets Malate Dehydrogenase 1 in non-small cell lung carcinomas. *Biochem Biophys Res Commun* 2018;499:314-20.
 26. Wang J, Ding M, Zhu H, et al. Up-regulation of long noncoding RNA MINCR promotes non-small cell of lung cancer growth by negatively regulating miR-126/SLC7A5 axis. *Biochem Biophys Res Commun* 2019;508:780-4.
 27. Guo R, Hu T, Liu Y, et al. Long non-coding RNA PRNCR1 modulates non-small cell lung cancer cells proliferation, apoptosis, migration, invasion and EMT through PRNCR1/miR-126-5p/MTDH axis. *Biosci Rep* 2020;40:BSR20193153.
 28. Sutani T, Sakata T, Nakato R, et al. Condensin targets and reduces unwound DNA structures associated with transcription in mitotic chromosome condensation. *Nat Commun* 2015;6:7815.
 29. Kim DS, Choi YP, Kang S, et al. Panel of candidate biomarkers for renal cell carcinoma. *J Proteome Res* 2010;9:3710-9.
 30. Zang X, Qian C, Ruan Y, et al. Higher maternal embryonic leucine zipper kinase mRNA expression level is a poor prognostic factor in non-small-cell lung carcinoma patients. *Biomark Med* 2019;13:1349-61.
 31. Kato T, Daigo Y, Aragaki M, et al. Overexpression of KIAA0101 predicts poor prognosis in primary lung cancer patients. *Lung Cancer* 2012;75:110-8.
 32. Schneider MA, Christopoulos P, Muley T, et al. AURKA, DLGAP5, TPX2, KIF11 and CKAP5: Five specific mitosis-associated genes correlate with poor prognosis for non-small cell lung cancer patients. *Int J Oncol* 2017;50:365-72.
 33. Isakova JT, Vinnikov D, Kipen VN, et al. Gene-to-gene interactions and the association of TP53, XRCC1, TNFalpha, HMMR, MDM2 and PALB2 with breast cancer in Kyrgyz females. *Breast Cancer* 2020;27:938-46.
 34. Ogata T, Teshima T, Inaoka M, et al. Carbon ion irradiation suppresses metastatic potential of human non-small cell lung cancer A549 cells through the phosphatidylinositol-3-kinase/Akt signaling pathway. *J Radiat Res* 2011;52:374-9.

(English Language Editor: J. Reynolds)

Cite this article as: Jiao Z, Yu A, He X, Xuan Y, Zhang H, Wang G, Shi M, Wang T. Bioinformatics analysis to determine the prognostic value and prospective pathway signaling of miR-126 in non-small cell lung cancer. *Ann Transl Med* 2020;8(24):1639. doi: 10.21037/atm-20-7520