

Prognostic values and prospective pathway signaling of miR-126 in non-small cell lung cancer: a study based on gene expression omnibus and bioinformatics analysis

Zichen Jiao

Nanjing drum tower hospital, medical school of nanjing university

Ao Yu

nanjing drum tower hospital, medical school of nanjing university

Xiaofeng He

nanjing drum tower hospital, medical school of nanjing university

Yulong Xuan

nanjing drum tower hospital, medical school of nanjing university

He Zhang

Nanjing drum tower hospital, medical school of nanjing university

Guojun Wang

jintan people's hospital, changzhou

Minke Shi

nanjing drum tower hospital, nanjing medical university

tao Wang (✉ wangtao_986@126.com)

Nanjing Drum Tower hospital, Medical school of nanjing university

Research

Keywords: NSCLC, MiR-126, Differentially Expressed Genes, Functional Factors

Posted Date: May 27th, 2020

DOI: <https://doi.org/10.21203/rs.3.rs-30704/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Objective MiRNAs are considered to be crucial for NSCLC's initiation and development. MiRNAs have been widely identified in NSCLC. However, the role of miR-126 in NSCLC has not been fully explained.

Methods miR-126 Expression in NSCLC was evaluated by analyzing the common data sets in Gene Expression Omnibus(GEO) database and reviewing former thesis papers. Three mRNA datasets, GSE18842, GSE19804 and GSE101929, from GEO to identify the differentially expressed genes (DEG). We prognosed the target genes of hsa-miR-126-5p using TargetScan and analyzed the gene overlap between the target genes of miR-126 and DEG in NSCLC. Subsequently, we analyzed Gene Ontology (GO) enrichment and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways. We used STRING and Cytoscape to construct a protein-protein interaction (PPI) network, and analyzed the influence of HUB gene on the prognosis of NSCLC.

Results A common pattern of mir-126 downregulation in NSCLC was identified in the literature review. A total of 187 DEGs were identified, both NSCLC-related and miR-126-related. Many DEGs are extendedly enriched in cell membranes, signal receptor binding, and biological regulation. Among the 10 main Hub genes analyzed by PPI, 4 HUB genes (NCAP-G, MELK, KIAA0101, TPX2) were obviously related to the poor recuperation of NSCLC patients. When these genes highly expressed, survival rate of NSCLC patients was low. Furthermore, we identified the recessive miR-126-related genes that may be involved in NSCLC, such as TPX2, HMMR, and ANLN through network analysis.

Conclusion this study suggests that mir-126 is radical for the biological processing of NSCLC.

1 Introduction

Lung cancer ranks first in global cases of malignant tumor, and the vast majority are non-small cell lung cancer (NSCLC). Due to its high degree of malignancy, there is currently no efficient treatment, and the fatality rate also ranks first in the world.[1] In recent years, the occurrence and fatality of lung cancer in our country have gradually increased [2]. It is because the pathogenesis of lung cancer remains unclear and its early symptoms are not typical, that most lung cancer patients are not diagnosed by histopathology until their middle or late stages of the disease, missing the best chance for early surgery. Even if they are diagnosed with lung cancer early enough to have the surgery, after tumor is totally removed, it is found during postoperative follow-up that the recurrence rate and metastasis rate are often as high as 1/3 [3]. With the in-depth study and investigation of driver-gene on lung cancer, we have a better understanding of the molecular biology of NSCLC, and the targeted therapy for driver genes has achieved remarkable results. Compared with traditional chemotherapy, targeted therapy has opened up a new therapeutic approach with high efficiency, low toxicity, accuracy and minimum adverse reactions in the field of lung cancer treatment, and thus improved the efficiency and safety of treatment towards NSCLC patients.[4, 5]. In particular, the maturity of next-generation sequencing (NGS) technology has further promoted the development of NSCLC driver gene detection and targeted therapy [6].

And thus we can conclude that the morbidity and fatality of NSCLC are so high. If we can find new targets related to the prognosis of NSCLC and clarify its mechanism, targeted treatment of NSCLC patients will greatly improve the clinical treatment effect of lung cancer as well as patients' quality of life, which would be essential in monitoring cancer recurrence and guiding rehabilitation.

MicroRNA (miRNA) is one of the mostly conserved non-coding small molecule RNAs. It is complete complementary or incomplete complementary to the 3'UTR, CDS or promoter region of mRNA, at the post-transcriptional stage of mRNA, to inhibits mRNA translation or directly degrades target mRNA, so as to regulate gene expression [7]. Since more miRNA family and its various regulatory functions are found, it is explicit that they are playing multiple roles in physiological and pathological processes such as cell differentiation, proliferation and apoptosis; meanwhile, they are an important factor to explain the initiation and development of various tumors.[8, 9]. More and more studies suggested that the abnormal expression of miR-126-5p (miR-126) plays multiple roles in the biological processes of various tumors., Studies suggest that MicroRNA-126 hinders the viability of colorectal cancer cells by hindering mTOR-induced apoptosis and autophagy [10]. Through the regulation of epigenetics we can use miRNA as a biomarker for the early diagnosis and treatment of malignant mesothelioma[11]. Kim et al[12] reported that ETV2 / ER71 regulates the production of FLK1 + cells from mouse embryonic stem cells through the miR-126-MAPK signaling pathway. On the other hand, some reports suggested that miR-126 can mediate the activation of STAT3 signal pathway, regulate the malignant biological behavior of NSCLC cells, and affect their proliferation, migration, cycle and susceptibility to apoptosis [13]. It has been explicit from previous studies that the expression level of miR-126 in NSCLC is numerously down-regulated and is crucial for NSCLC prognosis [14, 15].

Hence, on the basis of published data about the expression of miR-126 in NSCLC, this study aims to confirm possible molecular targets and explain the role of miR-126 in NSCLC by investigating miRNAs expression, GEO and literature review, as well as bioinformatics analysis.

2 Materials And Methods

2.1 Selection of GEO dataset

Gene chip map of non-small cell lung cancer (NSCLC) is obtained from GEO database (Gene Expression Omnibus,<http://www.ncbi.nlm.nih.gov/geo/>). The following key words were used for searching: (Lung) AND (cancer OR carcinoma OR tumor OR neoplasia OR neoplasm OR

Malignant OR malignancy) AND (microRNA OR miRNA OR noncoding

RNA OR ncRNA OR small RNA) Reports about the microarray data set of miR-126 expressed between non-small cell lung cancer and adjacent noncancerous lung tissue (ANLT) were included in our study.

2.2 The Research and Data Extraction of MiR-126 and Non-small Cell Lung Cancer in the Previous Studies

It was searched on PubMed, Web of Science, Embase and China national knowledge infrastructure (CNKI) (to date 1/4/2020), using key words: (MicroRNA OR miRNA OR noncoding RNA OR ncRNA OR small RNA) and (126 OR 126-5P) AND (Lung) and (cancer OR carcinoma OR tumor OR neoplasia OR neoplasm OR malignant OR malignancy).

Criteria for inclusion or non-inclusion were: (1) include studies of the expression of miR-126 in NSCLC; (2) exclude review, non-clinical studies, case reports, meta-analysis, and meeting summary; (3) exclude those with no control groups.

2.3 Gene ontology enrichment and target prediction analysis

The gene expression profiles of GSE18842, GSE19804 and GSE101929 were from Gene Expression Omnibus (GEO, <http://www.ncbi.nlm.nih.gov/geo>). The array data of GSE18842, GSE19804 and GSE101929 were composed by 46, 60 and 32 NSCLC samples and 45, 60 and 34 ANLT samples. All data were analyzed on the GPL570 platform: Affymetrix Human Genome U133 plus 2.0 (Affymetrix; Thermo Fisher Science, Inc., Waltham, MA, USA).

Limma software package (version 3.6.3) in R / BioManager was used to identify differentially expressed genes (DEG) between NSCLC and ANLT. By default, the adjusted P value (adj.PValue) uses Benjamini and Hochberg false discovery rate (FDR) methods to correct false positive results. $P \leq 0.05$ and $|\log_2(\text{FC})| \geq 1$ was set as the cut-off criterion. Based on the platform annotation file downloaded from the database, the probe data in the matrix file was converted into gene symbol.

The target gene of hsa-miR-126-5p (TG_miR-126-5p) was anticipated using the online website TargetScan (http://www.targetscan.org/vert_72/). Subsequently, the gene overlap between DEGs integrated in NSCLC was analyzed by bioinformatics software, and TG_miR-126-5p was anticipated. The gene overlap was analyzed through Gene Ontology (GO) and visualised by Binto plug-in of Cytoscape software (version 3.7.2). David database was applied to analyze the enrichment of Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway. $\text{FDR} \leq 0.05$ was set as the criterion. (STRING, version 11.0, <https://string-db.org/>) was used to build Protein-protein interaction (PPI) network of genes, whose interactions were also visualized by using Cytoscape software (version 3.7.2)'s CytoHubba plug-in. The confidence score $C \geq 0.7$ was set as the cut-off criterion. Then, we conducted the molecular complex detection (MCODE), and selected the PPI network modules with degree cutoff = 2, node score cutoff = 0.2, k-core = 2, and max. Depth = 100.

2.4 Survival analysis

Online tool Kaplan-Meier plotter (KM plotter, www.kmplot.com) will assess the influence of 21 cancers on survival rate, including some of the most enumerate ones: breast (n = 6,234), ovarian (n = 2,190), lung (n = 3,452),

And gastric (n = 1,440) cancer. Based on the median expression level of specific genes, we divided NSCLC patients into high expression and low expression groups. And we analyzed NSCLC patients' living situation by Kaplan-Meier method. Survival analysis was conducted on the general population; smoking population and non-smoking population. Then this study included the specific genes that showed statistical importance in the survival analysis of smoking population and non-smoking population. It was calculated and displayed the hazard ratio (HR) and its 95% confidence interval. Data set GSE102287 was applied to proof the correlation between miR-126 and DEGs. 62 samples in total were included in this analysis.

2.5 Statistical analysis

Data are displayed as mean \pm standard deviation (SD). The discrepancy between the two groups was investigated through two independent sample T tests. We also analyzed the correlation between miR-126 level and NSCLC by standardized mean difference (SMD) analysis by using Stata 15.0 statistical software. We used Mantel-Haenszel formula (fixed effect model) or Der Simonian-Laird formula (random effect model) to combine and analyze different GEO data sets. When the I^2 statistic matters ($P \leq 0.05$ or $I^2 \geq 50\%$), we adopt the random effect model and otherwise the fixed effect model. Spearman rank correlation was used to analyze the correlation between DEGS expression and miR-126 levels. When $P \geq 0.05$, it was assumed to statistically matter.

3 Results

3.1 miR-126 expression in OC based on GEO

Based on the GEO dataset (Fig. 1), the expression of miR-126 was evaluated in a series of NSCLC and ANLT. This study collected 4 GEO data sets (GSE102286, GSE29248, GSE63805 and GSE27705) in total. In the data set GSE102286, GSE29248, GSE27705, the expression level of miR-126 in NSCLC tissue was obviously lower than that of ANLT ($P < 0.01$); while in the data set GSE63805, the expression level of miR-126 in NSCLC tissue, compared with that of ANLT group, showed no statistically important discrepancy. The research characteristics based on the GEO dataset was enumerated in Table 1 and Fig. 2. In accordance with all included GEO data sets, the difference between non-small cell lung cancer and the normal control group was statistically important (SMD = -1.69, 95% CI: -2.97 ~ -0.41, $p < 0.01$). The results of forest map were as shown in Fig. 3.

Table 1
 Characteristics of studies based on GEO dataset

Study	NSCLC			Normal Control			t	p
	Mean	SD	n	Mean	SD	n		
GSE102286	12.59	0.72	90	14.34	0.71	89	16.29	0.000
GSE29248	13.18	0.23	6	13.58	0.19	6	3.31	0.008
GSE63805	13.79	1.19	31	14.15	1.09	31	1.24	0.220
GSE27705	0.45	0.22	20	2.07	1.27	10	4.00	0.003
Total	SMD=-1.69(95%CI:-2.97~-0.41), $p = 0.009$, $I^2 = 93.3%$, $p = 0.000$							
Note: NSCLC: Non-small cell lung cancer; SD: Standard deviation; SMD: standard mean difference								

3.2 Expression Profile of miR-126 in NSCLC and ANLT in Literature

Next, we investigated the expression of miR-126 in non-small cell lung cancer based on previous data. As shown in Fig. 4, we selected six studies from the literature that meet the selection criteria [14–19]. All studies showed that the expression level of miR-126 in NSCLC tissues was obviously lower than that in normal control group (Table 2).

Table 2
Overview of the studies selected from 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

Note: ANLT=adjacent noncancerous lung tissues;BLT=benign lung tissue=NSCLC=Non-small cell lung cancer=qRT-PCR=Quantitative Real-time Polymerase Chain Reaction

3.3 MiR-126 Prediction and Bioinformatics Analysis, Data Preprocessing and DEGS Screening

1829, 804, and 1921 DEGs were identified from the GSE18842, GSE19804, and GSE101929 data sets, respectively; 595 common DEGs were screened out in these three data sets by Venny 2.1.0 (<https://bioinfogp.cnb.csic.es/tools/venny/>) (Fig. 5, Fig. 6). Then, based on TargetScan, we anticipated 5591 TG_hsa-miR-126-5p, and 187 of which were proofed in 595 common DEGs. The data in GSE101929 showed that there were 46 up-regulated hsa-miR-126-5p related genes and 141 down-regulated hsa-miR-126-5p related genes in NSCLC tissues compared with that in normal control group (Fig. 6). Among them, the top ten up-regulated and down-regulated hsa-miR-126-5p related genes in data set GSE101929 were shown in Table 3.

Table 3

Top 10 hsa-miR-126-5p-related differentially expressed genes in NSCLC compared with normal lung tissues according to data from Gene Expression Omnibus (GSE101929)

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

3.4 Functional analysis of miR-126-related DEGs in NSCLC

Bingo plug-in in Cytoscape was used for Go function analysis. David was used for KEGG enrichment analysis. Analysis showed that many target genes were involved in biological processes such as cell

membrane, signal receptor binding, glycosaminoglycan binding, biological regulation, and stimulus response (Fig. 7, Table 4). In addition, among these potential target genes, there was an over expression of the KEGG pathway, that is the protein digestion and absorption interaction pathway (Table 4).

Table 4
Functional and pathway enrichment analysis of hsa-miR-126-5p-related differentially expressed genes in NSCLC

Term	Description	Count	FDR
Biological processes			
GO:0009987	cellular process	157	0.0408
GO:0065007	biological regulation	133	0.0276
GO:0050896	response to stimulus	101	0.0043
Molecular Function(Go)			
GO:0005102	signaling receptor binding	30	0.0221
GO:0005539	glycosaminoglycan binding	11	0.0073
GO:1901681	sulfur compound binding	10	0.0221
Cellular Component			
GO:0016020	membrane	107	0.0024
GO:0071944	cell periphery	88	1.38E-06
GO:0005886	plasma membrane	86	1.58E-06
KEGG pathway			
hsa04974	Protein digestion and absorption	7	0.0061
FDR: false discovery rate;KEGG=Kyoto Encyclopedia of Genes and Genomes			

3.5 PPI Network Construction and Modules Selection

The PPI network of miR-126-related DEG is composed of 187 nodes and 281 edges, including 46 up-regulated genes and 141 down-regulated genes (Fig. 8). Setting Degree as the ranking criterion, a total of Top 10 genes were selected as HUB genes, and there was a close correlation between HUB genes (Fig. 9A). An important module was obtained from the PPI network of mir-126-related DEG using MCODE, including 12 nodes and 64 edges (Fig. 9B).

3.6 Survival Analysis

The prognostic value of the top 10 HUB genes (CCNB1, KIAA0101, BUB1B, TPX2, NUF2, NCAPG, MELK, KIF15, HMMR, ANLN) in the PPI network were evaluated on the website <https://kmplot.com/analysis/>. The

overall survival of NSCLC patients was analyzed according to the high and low expression of each HUB gene. The results showed that all the ten genes were related to the prognosis of NSCLC in the overall survival analysis of population. However, the results were inconsistent when we divided the population into smoking and non-smoking groups. In the NSCLC survival analysis, genes with statistical differences between smokers and non-smokers were included in this study, and only four genes were up to criterion. Results showed that high expression of NCAPG (HR = 1.59, 95%CI :1.40 ~ 1.80, P < 0.01), high expression of MELK (HR = 1.6, 95%CI :1.41 ~ 1.82, P < 0.01), high expression of KIAA0101 (HR = 1.56, 95%CI :1.37 ~ 1.77, P < 0.01) and high expression of TPX2 (HR = 1.86, 95%CI:1.64 ~ 2.12, P < 0.01) were associated with the overall survival rate decreasing in patients with NSCLC (Fig. 10). In PPI network, the Top 10 Hub genes were negatively correlated with miR-126 to some extent, especially the genes TPX2, HMMR and ANLN, which play a potential role in NSCLC through interaction with mir-126 (Table 5).

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

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

4 Discussion

In the current study, we confirmed the abnormal expression of miR-126 related to NSCLC by comparing the expression profiles of miRNA in NSCLC tissues and normal control lung tissues based on the data from the GEO dataset and published studies. In addition, through GO analysis, KEGG analysis, protein-protein interaction (PPI) network and kaplan-meier plotter, we discovered and analyzed new markers and potential targets of miR-126 involved in the regulation of key biological processes in NSCLC.

So far, there have been few studies on the characteristics of miR-126 in NSCLC, but different studies have shown that the expression of miR-126 in NSCLC patients is down-regulated. A study of Wang et al[20]in

2015 showed that the level of miR-126 in the blood of patients with early stage non-small cell lung cancer was lower than that of healthy people. Zhu et al [21] and Shang et al [22] reached the same conclusion. Interestingly, in addition to being in the blood, the level of miR-126 in the sputum of lung cancer patients also decreased compared with the healthy control group. Their studies all suggested that miR-126 can be used as a marker for early diagnosis of non-small cell lung cancer. The expression of miR-126 in patients with non-small cell lung cancer decreased not only in body fluids, but the expression of miR-126 also lowered in non-small cell lung cancer tissues, comparing with adjacent tissues [14, 15]. And this decrease is significantly associated with the prognosis of patients with NSCLC [16, 17]. According to the included GEO dataset, three out of the four GSE datasets showed that miR-126 significantly decreased in NSCLC, and a meta-analysis of the random effect model showed that the combined SMD also had statistical importance. Therefore, according to the current research results, it will be of great significance for the treatment and prognosis of NSCLC patients to further study the mechanism of miR-126 in NSCLC, investigate its target genes and study its potential therapeutic targets.

MiR-126 is one of the most important miRNAs in the regulation of NSCLC. A meta-analysis by Weng et al [23] suggested that high expression of miR-126 was positive factor hopefully to improve the overall survival in patients with non-small cell lung cancer. MiR-126-3p can hamper the growth, migration and invasion of non-small cell lung cancer by targeting CCR1 [24]. Lima et al [25] reported that the expression of miR-126-5p hampered the enzyme activity of MDH1 and mitochondrial respiration in NSCLC cells, leading to cell death. The up-regulation of long-chain non-coding RNA MINCR promoted the growth of non-small cells in lung cancer by negatively regulating the miR-126 / SLC7A5 axis [26]. Guo et al [27]'s research showed that long-chain non-coding RNA PRNCR1 regulates the proliferation, apoptosis, migration, and invasion of non-small cell lung cancer cells through the PRNCR1 / miR-126-5p / MTDH axis. In view of the current situation, it is necessary to further clarify the molecular mechanisms and clinical value associated with abnormal expression of miR-126 in NSCLC.

Through PPI network analysis, we found ten Hub genes with high correlation. Through the survival analysis by online tools, we found that these ten Hub genes were associated with the prognosis of NSCLC. However, considering the important role of smoking in lung cancer, we conducted a separate prognostic analysis of smoking population and non-smoking population, and found that after stratified analysis, some genes became no longer significant, and only four genes' expression was significantly related to the prognosis of NSCLC, no matter the population smokes or not. These four genes are NCAPG, MELK, KIAA0101 and TPX2. Their high expression was associated with poor prognosis in NSCLC patients. This is also consistent with the negative regulation of miR-126. NCAPG (non-SMC condensin I complex, subunit G) is a subunit of the agglutinin complex, which is responsible for the condensation and stabilization of chromosomes during mitosis and meiosis [28]. Kim et al [29] reported that NCAPG can be used as a candidate biomarker for renal cell carcinoma. The maternal embryo leucine zipper kinase (MELK) belongs to the CAMK serine / threonine protein kinase superfamily and has maximum activity during mitosis. Zhang et al's [30] research showed that NSCLC patients with high MELK expression had a poor prognosis. This is consistent with the conclusion of this study. KIAA0101 mainly plays a role in the liver, and there are few studies with other organs. Kato et al [31] reported that the overexpression of

KIAA0101 predicted a poor prognosis in patients with primary lung cancer. This is consistent with the result of this study. Schneider et al[32] reported that TPX2 is associated with poor prognosis in patients with non-small cell lung cancer. It can be concluded that these genes are expected to become new prognostic markers for NSCLC.

In the current study, we found that new candidate target genes of miR-126 are involved in the regulation of key biological processes of NSCLC, such as TPX2, HMMR and ANLN. The HMMR gene has been proved to be a polyhedral molecule. First, HMMR has the role of promoting tumors. Its product hyaluronic acid-mediated cell migration receptor (RHAMM) has been proved to be a cancer promoter and may become a tumor cell migration promoting factor, whose expression promotes the occurrence and development of tumors, and is related to the pathological stage of tumors [33]. Some studies have shown that carbon ion irradiation can inhibit the expression of ANLN (ANLN) gene in human lung adenocarcinoma cells, which is regulated by the activation of phosphatidylinositol-3-kinase (PI3K) / Akt signal pathway associated with metastasis[34]. Our correlation analysis also showed that miR-126 was negatively correlated with genes TPX2, HMMR and ANLN in non-small cell lung cancer, and the difference was statistically significant, indicating that they play a very important role in the biological regulation of NSCLC.

In conclusion, the results of this study suggest that miR-126 plays an important role in NSCLC biology. However, further in vivo and in vitro experiments are needed to further study its pathogenesis and thus to clarify the role of the molecular network regulated by miR-126 in NSCLC.

Abbreviations

NSCLC: Non-small cell lung cancer ; miR-126: hsa-miR-126-5p; GEO: Gene Expression Omnibus; ANLT: adjacent noncancerous lung tissue ; DEG: differentially expressed genes ; GO: Gene Ontology ; KEGG: Kyoto Encyclopedia of Genes and Genomes ; PPI: protein-protein interaction ; NGS: next-generation sequencing ; miRNA: MicroRNA ; SD: standard deviation ; SMD: standardized mean difference .

Declarations

Ethics approval and consent to participate

Ethical approval was not needed because this is a Bioinformatics analysis.

Consent for publication

Not applicable.

Availability of data and material

The datasets used and/or analysed during the current study are available from the corresponding author on reasonable request.

Competing interests

All the authors declare that they have no conflict of interest.

Funding

Our work is supported by Natural Science Foundation of Jiangsu Province, Grant/Award Number: BK20140601

Author's contributions

All authors contributed to the study design; all authors collected the data and performed the data analysis; all authors prepared the manuscript.

Acknowledgements

None.

References

1. Siegel RL, Miller KD, Jemal A. Cancer statistics. 2018. *Ca A Cancer Journal for Clinicians* 2018;68:7–30.
2. Rimm EB, Giovannucci EL, Stampfer MJ, Colditz GA, Litin LB, Willett WC. Reproducibility and Validity of an Expanded Self-Administered Semiquantitative Food Frequency Questionnaire among Male Health Professionals. *AM J EPIDEMIOLOG* 1992;10.
3. Port JL, Parashar B, Osakwe N, Et Al. A Propensity-Matched Analysis of Wedge Resection and Stereotactic Body Radiotherapy for Early Stage Lung Cancer. *ANN THORAC SURG*. 2014;98:1152–9.
4. Sacco C, Maione P, Guida P. C, Et Al. The Combination of New Immunotherapy and Radiotherapy: A New Potential Treatment for Locally Advanced Non-Small Cell Lung Cancer. *Current Clinical Pharmacology* 2017;12:4.
5. Bagley SJ, Talento S, Mitra N, Meropol NJ, Cohen RB, Langer CJ, 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, Despain D, Choquette AK, Spira AI, et al. Cell-Free DNA Next-Generation Sequencing Prediction of Response and Resistance to Third-Generation EGFR Inhibitor. *CLIN LUNG CANCER*. 2018;19:518–30.
7. Lujambio A, Lowe SW. The microcosmos of cancer. *NATURE*. 2012;482:347–55.
8. Fan J, Xu W, Nan S, Chang M, Zhang Y. 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, Georges S, Baud'Huin M, Lamoureux F, et al. Implication of the p53-Related miR-34c, -125b, and -203 in the Osteoblastic Differentiation and the Malignant Transformation of Bone Sarcomas. *CELLS-BASEL* 2020;9.
10. Wei L, Chen Z, Cheng N, Li X, Chen J, Wu D, et al. MicroRNA-126 Inhibit Viability of Colorectal Cancer Cell by Repressing mTOR Induced Apoptosis and Autophagy. *ONCOTARGETS THER.* 2020;13:2459–68.
11. Tomasetti M, Gaetani S, Monaco F, Neuzil J, Santarelli L. 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, Choi HS, Dwivedi B, Rupji M, 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, Yu X. 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, Jabboue FJ, Brait M, Izumchenko E, 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. Switlik W, Karbownik MS, Suwalski M, Kozak J, Szemraj J. 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 JY, Lan HT, Huang XB, Liu BY, Tong Y. 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, Su CQ, Dong HF. 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, Zheng YG. 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 PATHO.* 2015;8:14983–8.
20. Wang P, Yang D, Zhang H, Wei X, Ma T, Cheng Z, et al. Early Detection of Lung Cancer in Serum by a Panel of MicroRNA Biomarkers. *CLIN LUNG CANCER.* 2015;16:313–9.
21. Zhu W, Zhou K, Zha Y, Chen D, He J, Ma H, 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:e153046.
22. Shang AQ, Xie YN, Wang J, Sun L, Wei J, Lu WY, et al. Predictive 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, Xu H, Lei L, Chen C, 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, Cheng ZM, Yu JL, Zhou 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, Hao Y, Eriksson M, Hydbring P, et al. miR-126-5p targets Malate Dehydrogenase 1 in non-small cell lung carcinomas. *BIOCHEM BIOPH RES CO.* 2018;499:314–20.
26. Wang J, Ding M, Zhu H, Cao Y, Zhao W. 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, He Y, Cao Y. 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.
28. Sutani T, Sakata T, Nakato R, Masuda K, Ishibashi M, Yamashita D, 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, Gao MQ, Kim B, Park HR, et al. Panel of candidate biomarkers for renal cell carcinoma. *J PROTEOME RES.* 2010;9:3710–9.
30. Zang X, Qian C, Ruan Y, Xie J, Luo T, Xu B, 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, Ishikawa K, Sato M, Kaji M. Overexpression of KIAA0101 predicts poor prognosis in primary lung cancer patients. *LUNG CANCER.* 2012;75:110–8.
32. Schneider MA, Christopoulos P, Muley T, Warth A, Klingmueller U, Thomas M, 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, Talaibekova ET, Aldashev AA, Aldasheva NM, 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-TOKYO* 2020.
34. Ogata T, Teshima T, Inaoka M, Minami K, Tsuchiya T, Isono 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.

Figures

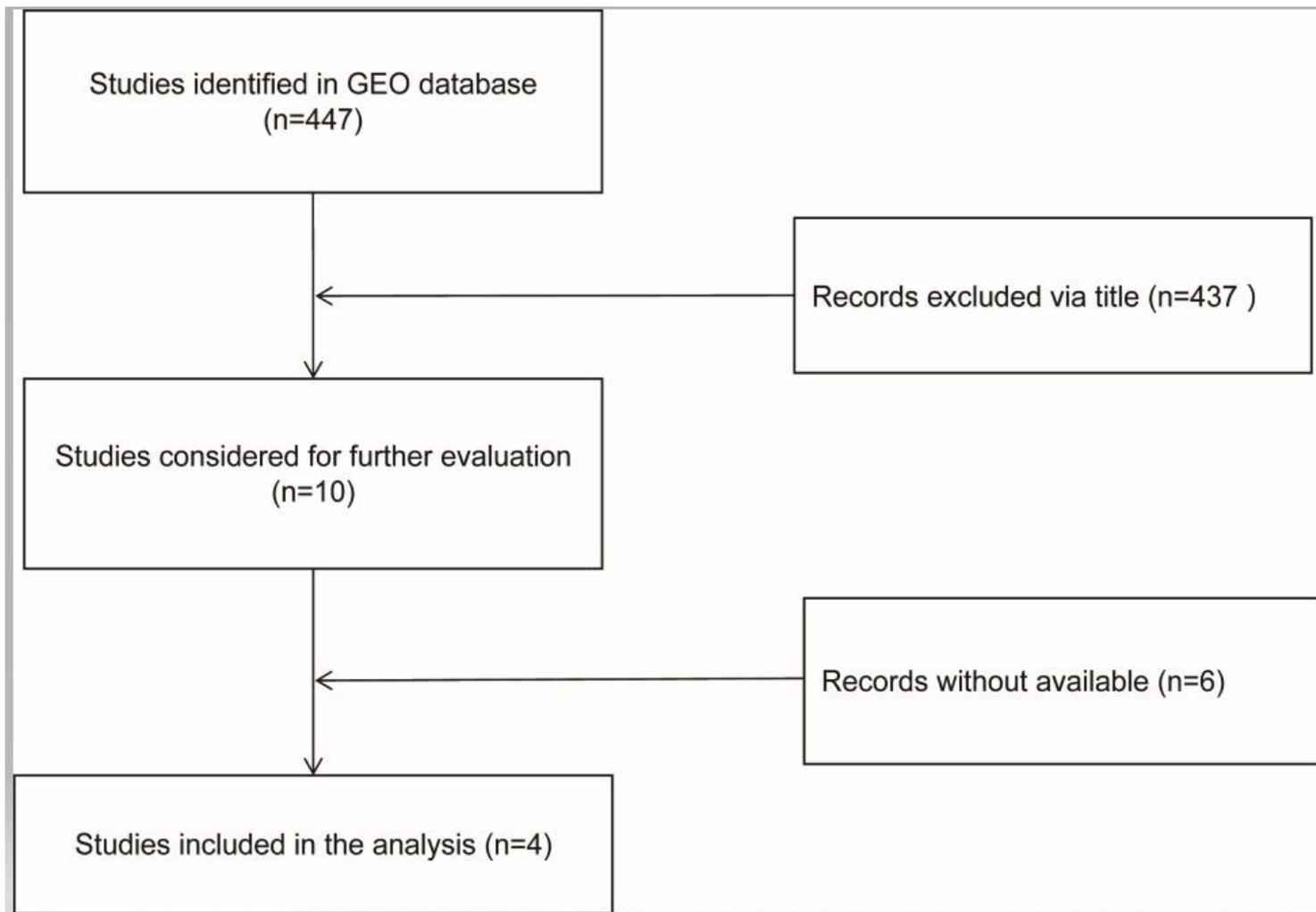


Figure 1

Flow diagram of study selection for GEO dataset

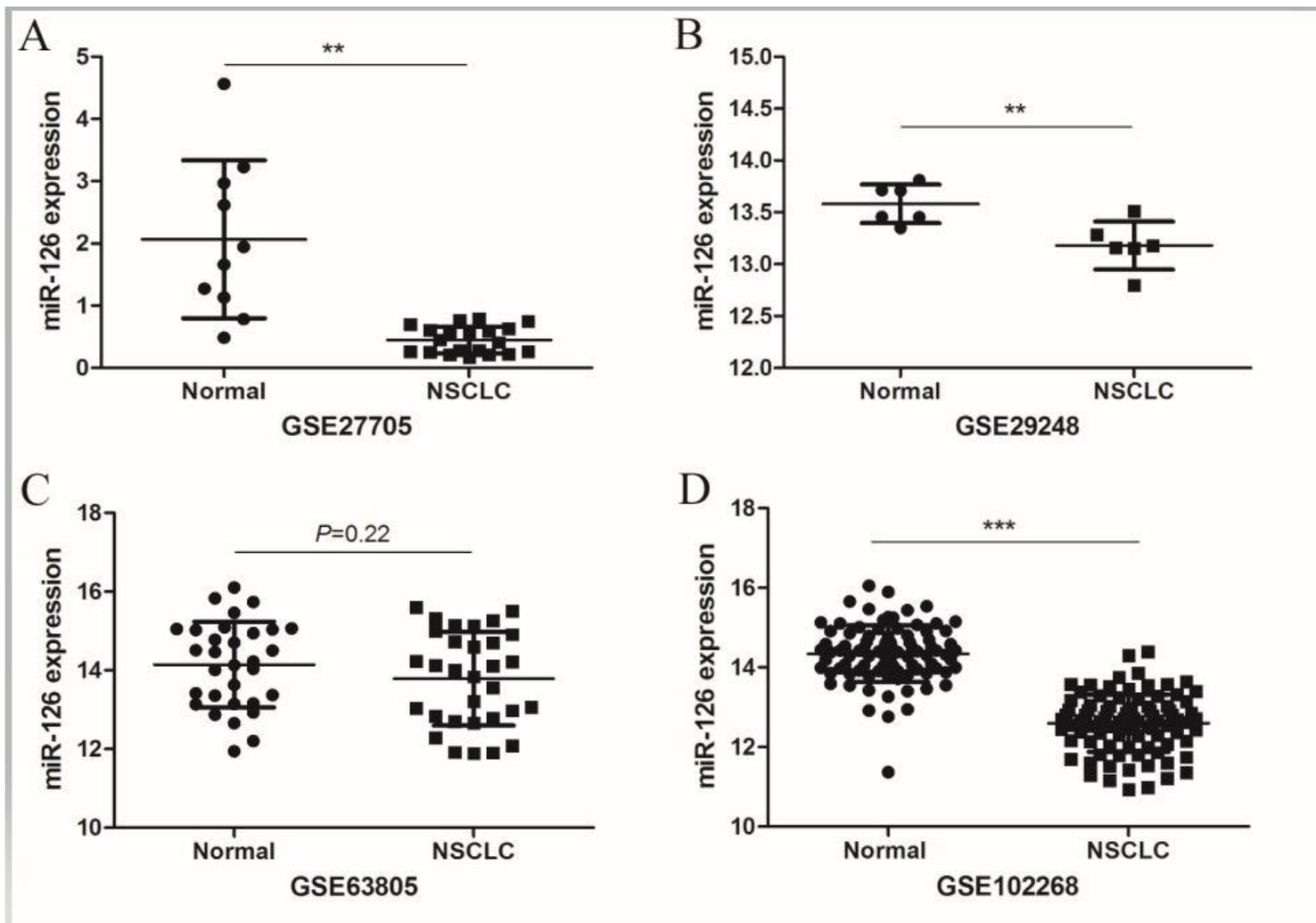


Figure 2

Expression of miR-126 in NSCLC and normal lung tissues in GEO datasets. NSCLC: Non-small cell lung cancer; Normal: normal lung tissue; miR-126: hsa-miR-126-5p

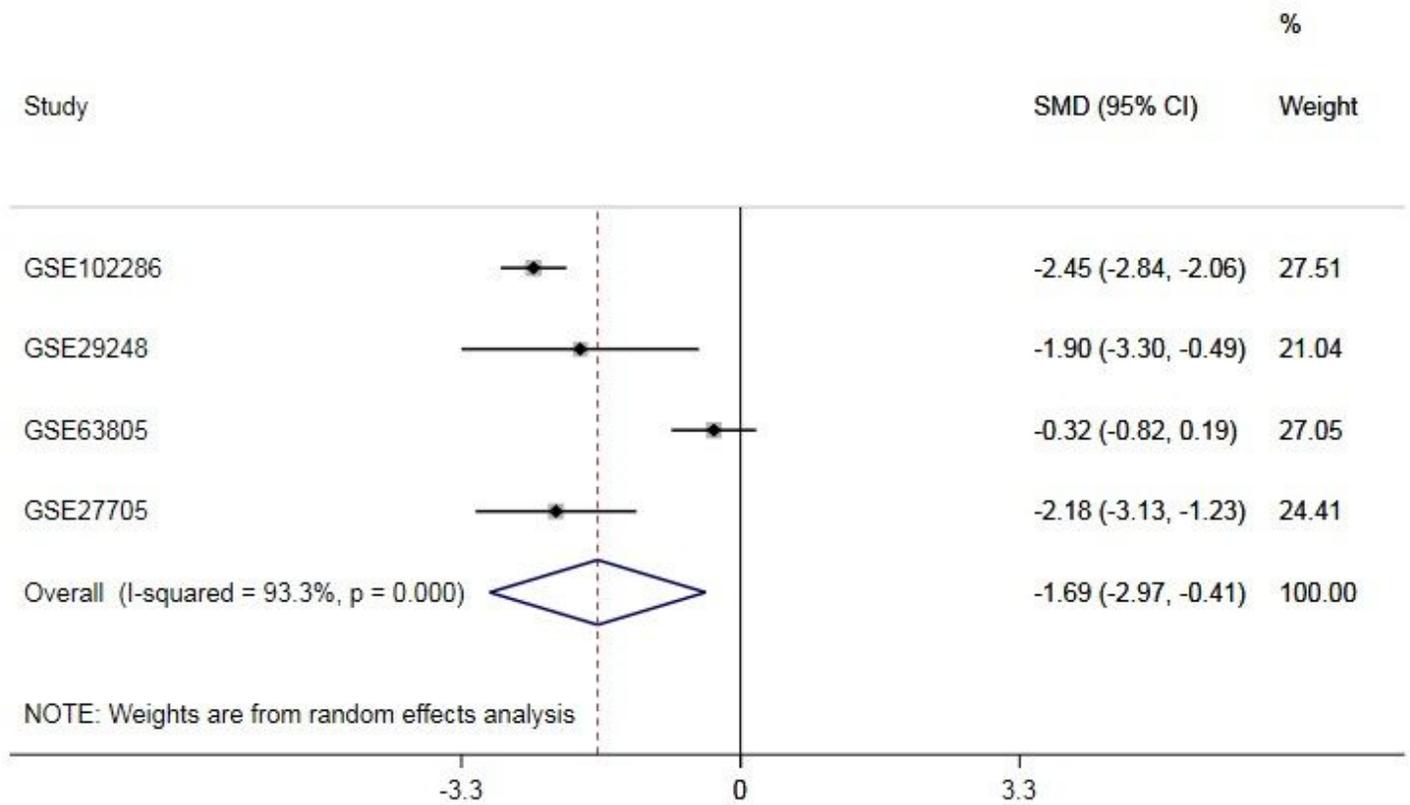


Figure 3

Forest plot of the combined SMD for hsa-miR-126-5p expression between NSCLC and normal lung tissue by the random effects models

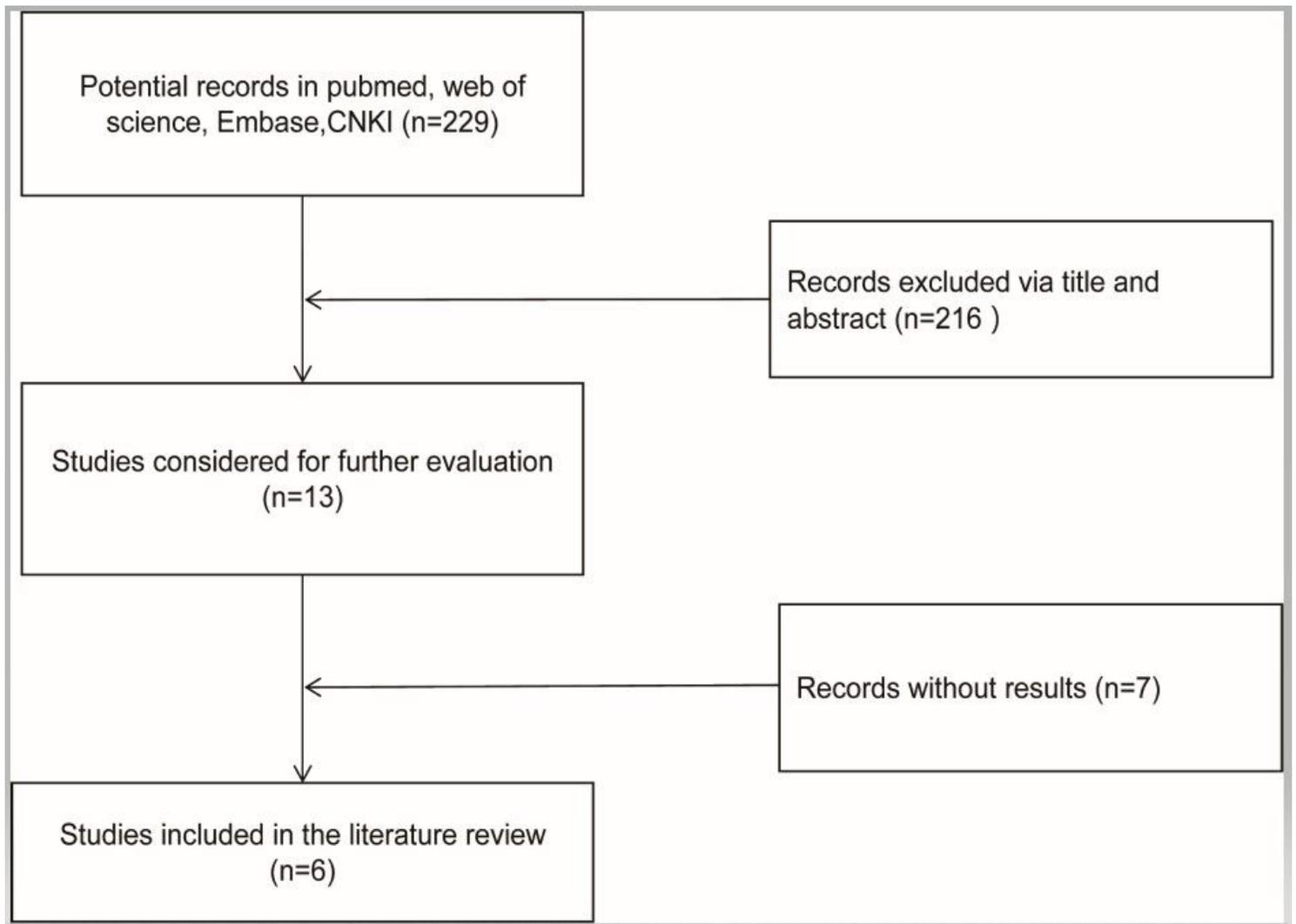


Figure 4

Flow chart of study selection for the literature review

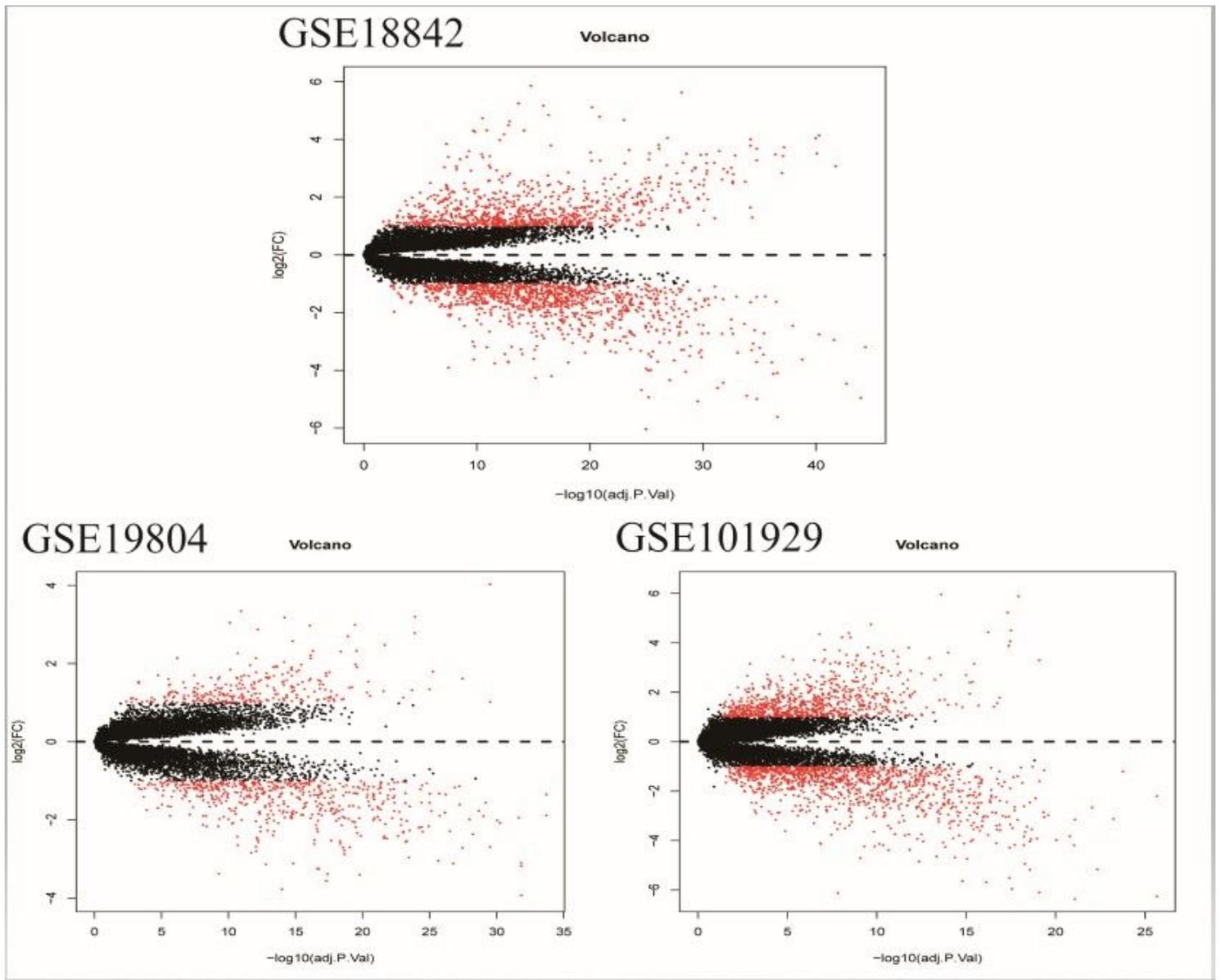


Figure 5

Volcano plot of detectable genome-wide mRNA profiles in NSCLC tissue and normal lung tissue samples from GSE18842, GSE19804, and GSE101929, respectively. Red plots represent aberrantly expressed mRNAs with $P < 0.05$ and $|\log_2(\text{FC})| > 1$. Red plots above indicate up-regulated genes, red plots below indicate down-regulated genes and black plots indicate normally expressed mRNAs. The x-axis is the $-\log_{10}$ of the adj.P.Value for each mRNA, representing the strength of the association. adj.P.Value, adjusted P value; FC, fold change

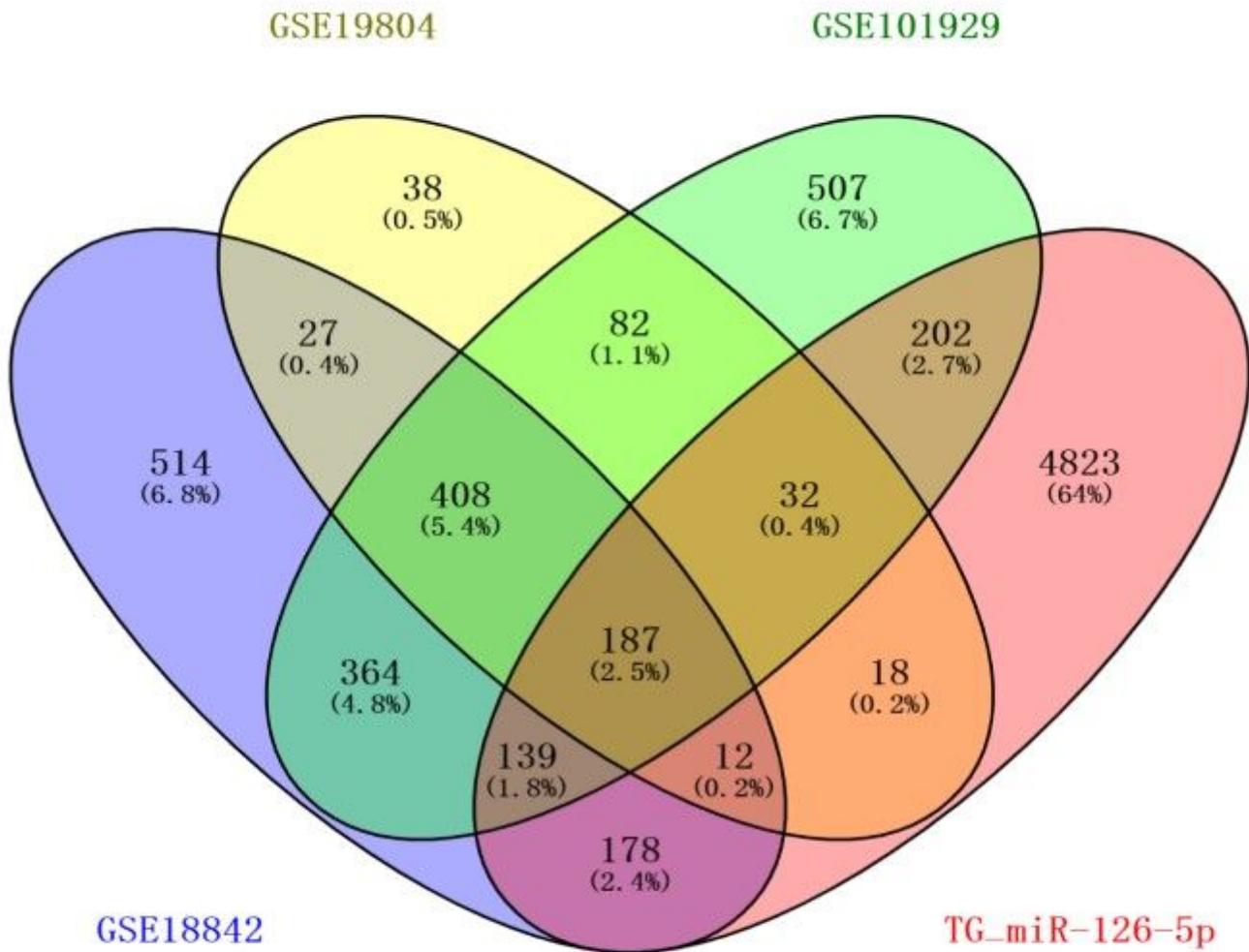


Figure 6

Venn plots of hsa-miR-126-5p-related differentially expressed genes from four datasets (GSE18842, GSE19804, GSE101929, and TG_miR-126-5p), the overlapping area corresponds to the commonly identified DEGs. DEGs: differentially expressed genes; TG_miR-126-5p, target genes of hsa-miRNA-126-5p.

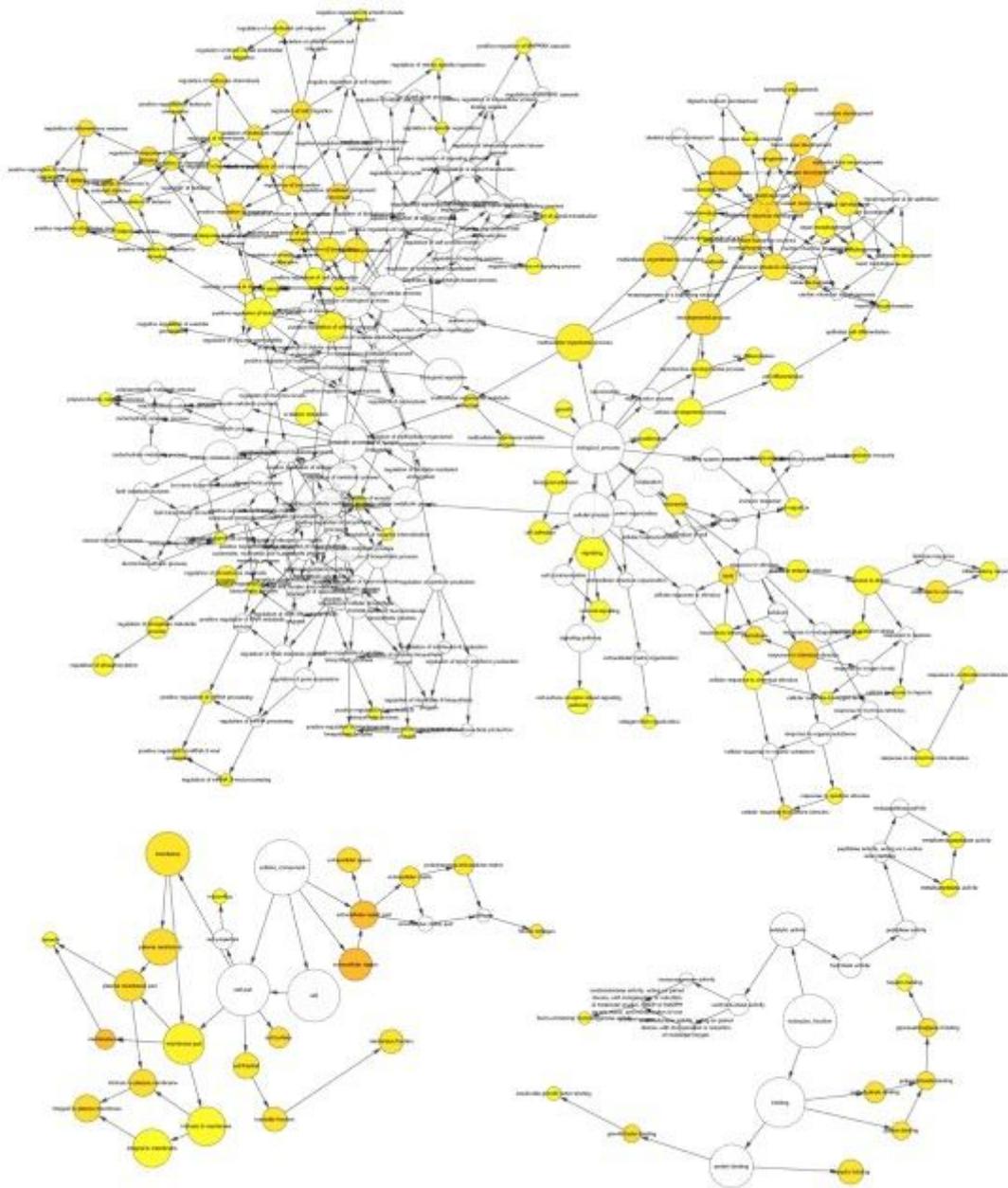


Figure 7

Results of GO enrichment Analysis of miR-126 NSCLC Target Gene. The yellow circle represents functional enrichment, and the larger the circle, the darker the color, the more genes are enriched in this pathway. The connecting lines represent the association between gene and gene.

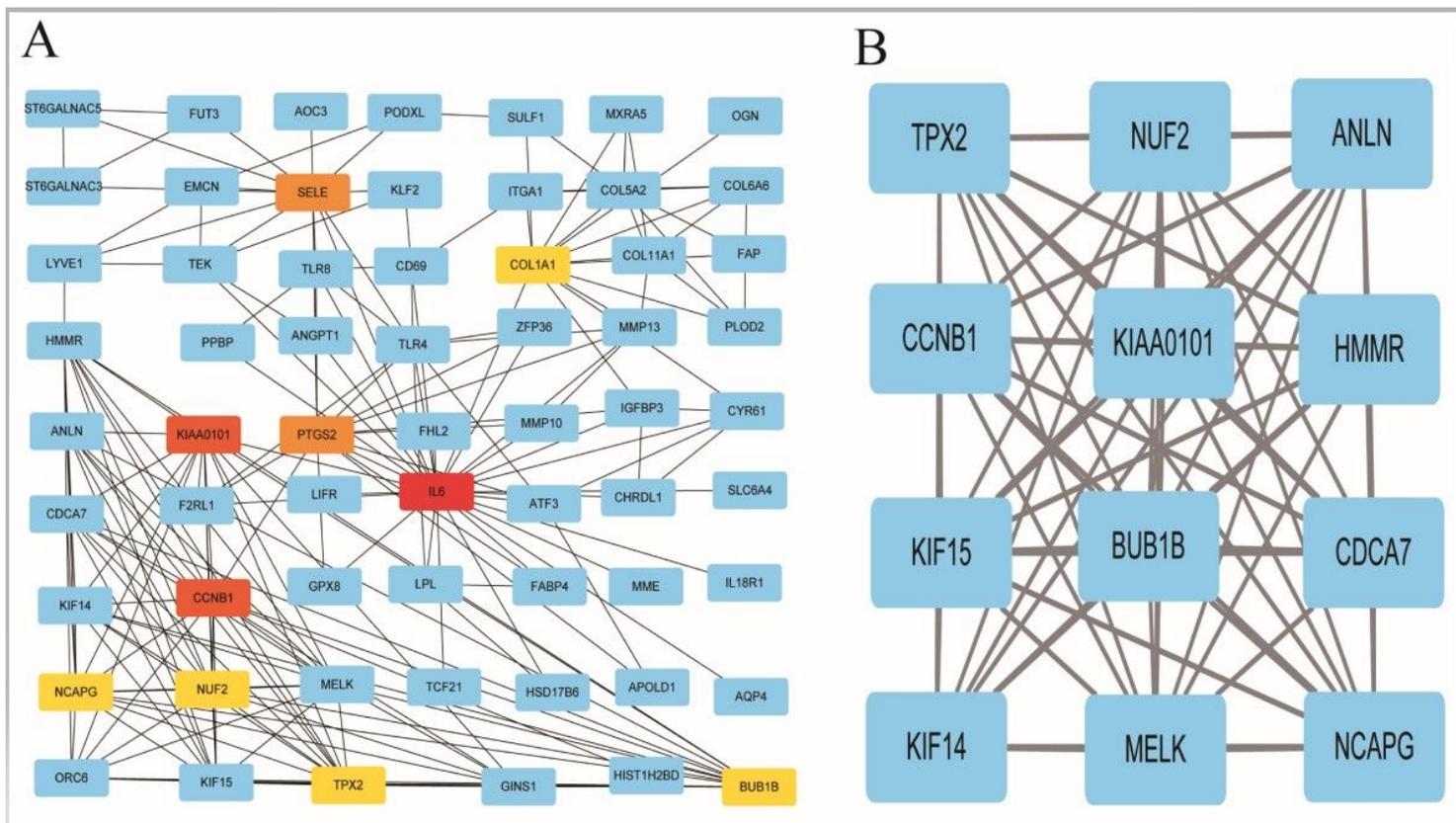


Figure 9

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 protein-protein interaction network of hsa-miR-126-5p-related DEGs. Red and yellow represent the key Hub genes of Tp10. 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.

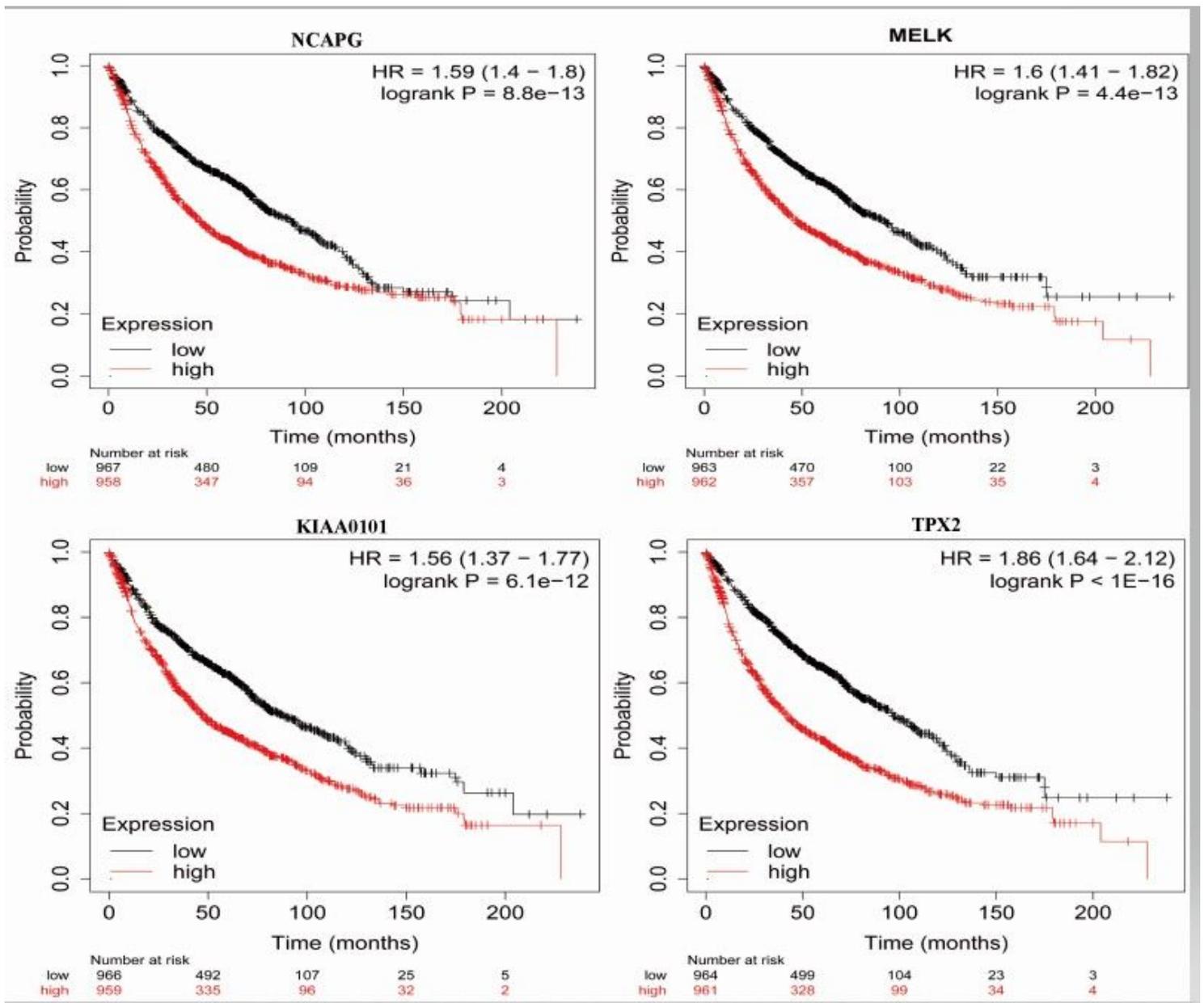


Figure 10

Overall survival analysis of NCAPG , MELK ,KIAA0101 and TPX2 expression with prognosis of NSCLC patients. The patients with NSCLC were divided into two groups (high vs. low), according to the median expression level.