ORIGINAL ARTICLE e-ISSN: 2349-0659 p-ISSN; 2350-0964

Identification of miRNA Biomarkers for Non-small Cell Lung Cancer using Microarray Datasets and Bioinformatics Analysis

Imteyaz Ahmad Khan, Srikant Sharma*

Abstract

Non-small cell lung cancer (NSCLC) is the most common type of lung cancer, which accounts for about 81% percent of all lung cancer cases. Despite advances in NSCLC diagnosis and treatment, the mortality rate is still high. Patients with NSCLC have a 5-year survival rate of only 21%. NSCLC is currently diagnosed at an advanced stage when cancer has already metastasized. Hence, it is essential to elucidate the molecular mechanisms associated with NSCLC pathogenesis and identify early diagnostic or predictive biomarkers. Circulating miRNAs might act as non-invasive blood-based biomarkers for NSCLC diagnosis and prognosis. In this study, we used previously published microarray datasets from Gene Expression Omnibus database and identified 12 down-regulated miRNAs in tissue and blood of NSCLC. Further analysis identified three miRNAs that could serve as biomarkers in the diagnosis of NSCLC: miR-140-3p, miR-29c, and miR-199a. The functional enrichment analysis of the candidate microRNA's target genes disclosed various overrepresented pathways associated with malignancy progression. Seven target genes are identified as hub genes of the PPI network and hold strong predictive power. In addition, the three-gene combination (IL6, SNAI1, and CDK6) shows a hazard ratio of more than one (hr = 1.5) and *P*-value <0.002. Since the expression levels of these three miRNAs were significantly decreased in both tissue and blood, measuring miRNA expression in the blood also gives information on its expression in tissue. Therefore, these three microRNAs could be used as NSCLC diagnostic and prognostic biomarkers.

Keywords: Biomarkers, Gene expression omnibus, miRNA, Non-small cell lung cancer, Protein-protein interaction network

Asian Pac. J. Health Sci., (2022); DOI: 10.21276/apjhs.2022.9.4.28

Introduction

Lung cancer is the most common cause of cancer-related deaths globally, with 1.8 million deaths annually.[1] It is the most commonly diagnosed malignancy in males, followed by colorectal, liver, and esophageal cancers; while in females, it is the most frequent cause of cancer-related deaths after breast cancer.[1] Based on histology, lung cancers are classified into two major groups: Non-small cell lung cancer (NSCLC) and small cell lung cancer (SCLC).[2] NSCLC is the most common type of lung cancer, which accounts for about 81% of all lung cancer cases.[2] Furthermore, NSCLC is divided into three subtypes: Lung adenocarcinoma, lung squamous cell carcinoma, and large cell carcinoma.[2] Despite advances in NSCLC diagnosis and treatment, the mortality rate is still high. Patients with NSCLC have a 5-year survival rate of only 21%.[3] The lethality of NSCLC is often attributed to a lack of early diagnosis, metastasis, and the occurrence of drug resistance. NSCLC is currently diagnosed at an advanced stage when the cancer has already metastasized. Hence, it is essential to elucidate the molecular mechanisms associated with NSCLC pathogenesis and identify early diagnostic or predictive biomarkers.

MicroRNAs (miRNAs) represent a class of endogenous, small (20–24 nucleotides), and non-coding single-stranded RNAs, which specifically bind to the 3′ UTR of mRNA targets to inhibit post-transcriptional gene regulation[4,5] A single miRNA can regulate hundreds of different target genes and more than 50% of the miRNAs are associated with cancer initiation, progression, and metastasis.[6-8] Therefore, some miRNAs that are leaked by cancer cells and circulate in a stable form in the circulation (blood) can be used as early non-invasive biomarkers for different cancer types. In a previous study, several miRNAs differentially expressed between NSCLC and controls have been identified in tissue and blood samples.[9-11]

Department of Biotechnology, Shri Venkateshwara University, Gajraula, Uttar Pradesh, India

Corresponding Author: Dr. Srikant Sharma, Department of Biotechnology, Shri Venkateshwara University, Gajraula, Uttar Pradesh, India. Mobile: 9259451426. E-mail: shribioinfo@gmail.com

How to cite this article: Khan IA, Sharma S. Identification of miRNA Biomarkers for Non-small Cell Lung Cancer using Microarray Datasets and Bioinformatics Analysis. Asian Pac. J. Health Sci., 2022;9(4):142-149.

Source of support: Nil

Conflicts of interest: None.

Received: 03/02/2022 **Revised:** 17/03/2022 **Accepted:** 15/04/2022

Blood-based miRNAs can be profiled using different techniques, including next-generation sequencing, microarray, and real-time PCR. In a previous study, microarray-based miRNA expression detection has been used extensively to identify biomarkers for different types of cancer, such as pancreatic ductal adenocarcinoma, breast cancer, colon cancer, and lung cancer.[12-14] In a recent study, two serum miRNAs (miR-1228-3p and miR-181a-5p) showed a promising result for the early detection of lung cancer, indicating that these miRNAs might be used as non-invasive biomarkers for lung cancer.[15]

Studies have demonstrated significant interactions between gene alterations and cancer development in different types of tumors.[16] A recent study showed that high expression of Notch3 and CD44 was associated with rapid NSCLC progression.[17,18] A recent study conducted by Morris *et al.* demonstrated that the expression levels of the FPR1 gene in blood samples predict both NSCLC and SCLC.[19] In a recent study, miR-30d has been shown to be a tumor suppressor in the progression of NSCLC.[20] MiR-598 functions as tumor suppressor in NSCLC by targeting several oncogenic genes.[21] However, the exact mechanism of miRNAs

©2022 The Author(s). This is an open access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

and genes dysregulated in NSCLC pathogenesis is still not properly understood.[22]

Recently, gene expression microarray and high-throughput RNA-sequencing have been frequently used technology applied in cancer biomarker identification.[23] In this study, four previously published miRNA microarray datasets (GSE137140, GSE93300, GSE94536, and GSE53882) were used to identify differentially expressed miRNAs in NSCLC patients compared with that from normal healthy controls. MicroRNA target prediction (TargetScan, PicTar, and miRanda) and gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis were performed to identify possible pathways involved. Finally, to further validate the potential of the identified miRNAs as candidate biomarkers of NSCLC, we performed survival and receiver operating characteristic (ROC) analysis.

Methods

Dataset Selection

The Gene Expression Omnibus (GEO) is a freely accessible online database that contains high-throughput functional genomic data from various studies.[24,25] GEO datasets were searched for the keywords blood (serum or plasma), NSCLC, Human, and miRNA profiling through the NCBI website. Four miRNA microarray datasets, GSE137140, GSE93300, GSE94536, and GSE53882 were selected for analysis [Table 1].[26-28] The GSE137140 dataset contained serum miRNA profiling data consisting of 1566 NSCLC samples and 1774 control samples, while the other two datasets (GSE93300 and GSE94536) contained plasma miRNA profiling data. In addition, we selected tissue miRNA data from the GSE53882 dataset, consisting of 397 NSCLC tissue and 151 control samples.[29]

Identification of Differentially Expressed miRNAs

All the microarray datasets were analyzed for differentially expressed miRNAs (DEmiRNAs) in the R environment (R version 3.0, www.r-project.org) and the bioconductor limma package.[20] The statistical significance of the fold change was calculated for each miRNA by Student's *t*-test. The normal distribution of gene expression values was verified using the online tool GEO2R. Based on differential expression analysis, the DEmiRNAs were separated based on up-regulation or down-regulation. The cutoffs for differentially expressed miRNAs were set as absolute fold-change >2 and *P*-value of <0.05. To visualize the number of differently expressed miRNAs for each dataset, Venn diagrams were plotted.[21] GraphPad Prism software 4.0 (USA) was used to construct a heatmap.

Table 1: Summary of four datasets employed in this study

Dataset ID	*Platform*	*No. of NSCLC samples*	*No. of control samples*	*References*
GSE137140	GPL21263	1566	1774	[26]
GSE93300	GPL21576	9	4	[27]
GSE94536	GPL21576	6	3	[28]
GSE53882	GPL18130	397	151	[29]
Total	NA	1978	1932	

Predicting Target Genes

The target genes of the DE-miRNAs were predicted by miRDB (mirdb.org), TargetScan (www.targetscan.org), and mirTarBase (https://mirtarbase.cuhk.edu.cn/) prediction software.[30] The target gene list was analyzed and submitted to PANTHER for enrichment analysis of the significantly overrepresented GO biological processes and molecular function terms.[31] Fisher's exact test was used to determine statistical significance. *P*-value<0.05 was considered as the statistical significance level.

PPI and miRNA-mRNA Network

The candidate genes identified in this study were investigated in the STRING database, to construct a PPI network.[32] The confidence level of PPI was set to the highest (0.9). Then, cytoscape was used to create and visualize the graphs of the PPI network. In addition, the cytoscape was also used to screen hub DEGs with the node degree and clustering coefficient. Genes were considered as hub genes if their connection degrees are ≥20. In addition, the MCODE software was used within cytoscape to analyze PPI network modules with default cutoff criteria (degree cutoff ≥ 2, node score cutoff ≥2, K-core ≥ 2, and Max depth = 100). The STRING plugin was used for functional enrichment analysis on individual modules.

Analysis of Hub Genes

Kaplan–Meier survival curves were constructed to visualize two subtypes of NSCLC, adenocarcinoma (ADC), and squamous cell carcinoma (SCC).[33] The Kaplan–Meier plotter (http://www.kmplot.com) and an in silico web tool were used to merge the gene expression data from TCGA and the GEO database. Differentially expressed hub genes were identified using cytoscape. Finally, the hub genes were submitted to the Kaplan–Meier plotter (KM plotter). The patients' survival was analyzed using survival curves. For each gene, a hazard ratio (HR) with 95% confidence intervals and a log-rank *P*-value was calculated. Finally, the GEPIA software was used for the analysis of hub genes.[34]

Identification of Candidate miRNA Biomarkers for NSCLC

The Kaplan–Meier estimations were used to calculate the survival function and survival curves for the differentially expressed miRNAs. MiRNAs with statistically significant (log-rank p-value) prognostic potential were further analyzed using TargetScan, which predict potential miRNA targets. The log2 transformation was applied to identify differentially expressed miRNAs in lung adenocarcinoma and lung squamous cell carcinoma. GraphPad Prism 6.0 software (USA) was used to create the ROC curve.

Results

Identification of Differentially Expressed miRNAs

GSE137140, GSE93300, GSE94536, and GSE53882 were selected for analysis. Additional quantile normalization was performed for GSE93300 datasets. The GSE137140 dataset showed 280 up-regulated and 2200 down-regulated miRNAs. The GSE93300 dataset showed 2161 up-regulated and only 58 down-regulated miRNAs, The GSE94536 dataset contained 242 down-regulated miRNAs; however, no significantly up-regulated miRNAs were

found. Finally, the GSE53882 dataset showed 550 up-regulated and 564 down-regulated miRNAs.

The down-regulated differentially expressed miRNAs (DE-miRNAs) from each of the four datasets were entered into the web-based program VENNY 2.0 to generate 4-way Venn figures of overlapping differentially expressed miRNAs [Figure 1]. Thirteen differentially expressed miRNAs were common to all four datasets and 54 differentially expressed miRNAs were identified in three out of four microarray datasets. For further analysis, the 13 overlappings differentially expressed miRNAs were chosen, and their log fold-change values were calculated for each dataset [Figure 2].

Target Gene Prediction and Functional Enrichment Analysis of Differentially Expressed miRNAs

We identified 289 target genes for 12 differentially expressed miRNAs using TargetScan, mirTarBase, miRdb, and miRwalk. The

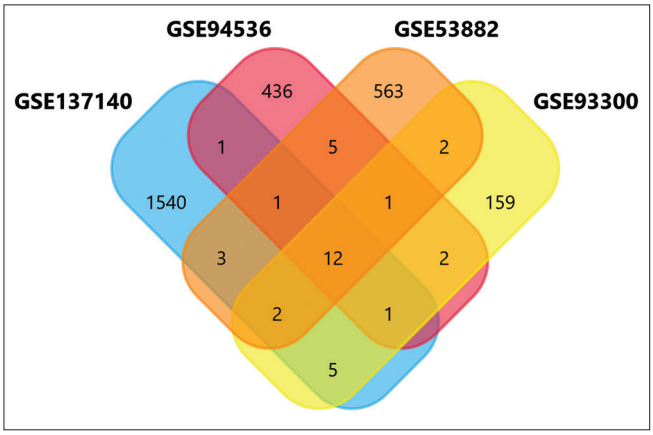

Figure 1: Four-set Venn diagram of the overlap of significantly differentially expressed miRNAs

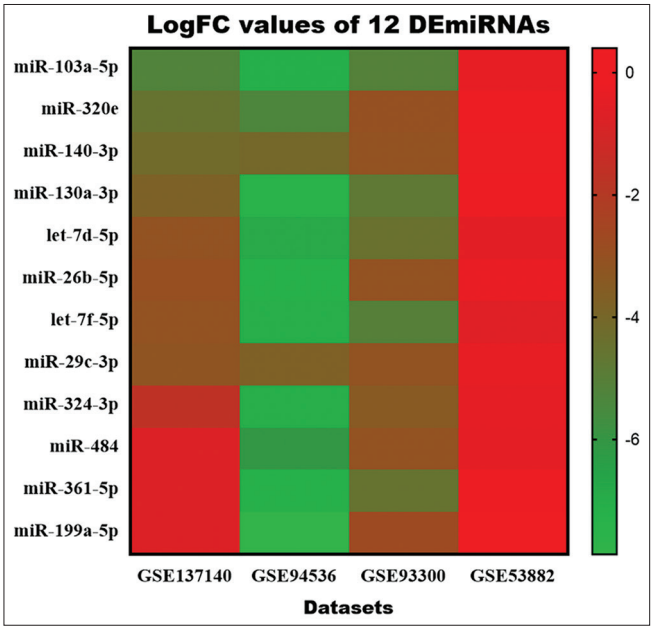

Figure 2: Heatmap shows the log fold change of the DEmiRNAs. The green colour indicates a low expression value and the red colour denotes a high expression value

expression of these genes is found to be over-expressed in NSCLCs because the miRNAs that regulate them are down-regulated. The predicted target genes of these differentially expressed miRNAs were analyzed using PANTHER and showed significantly enriched GO, various molecular functions, and biological processes. The top enriched biological processes and molecular functions are all involved in cancer progression, invasion, and metastasis.

Furthermore, it was observed that the most of the enriched biological processes were related to the epithelial to mesenchymal transition (EMT), phosphorylation of pathway-restricted SMAD proteins (17.11 times), heterochromatin assembly (12.21 times), cell growth and proliferation, negative regulation of gene silencing by miRNAs (11.45 times), and response to cholesterol. In addition, many other candidate miRNA target genes related to cellular components and molecular functions were also identified, which relate to transcription, protein kinase activity, transcription, and chromatin binding. In addition, PANTHER analysis applied to the GO cellular component terms showed that the RISC complex was enriched by 14.57 times.

PPI Network Analysis

We constructed a comprehensive protein-protein interaction network from 290 target genes using the STRING database. The result identified 678 interactions between nodes (average node degree of 3.89) [Figure 3]. The expected number of interactions was 498; the p-value for this high number of overlaps was <1.2E-15. Hub gene network connections were uploaded to cytoscape for visualization using the cytoscape plugin app MCODE or ClusterONE and network analyzer.

MCODE was used to identify the functional modules of the highly interconnected clusters of genes in the cytoscape network. Using MCODE algorithms, we identified eight modules. Among them, the top 2 modules are displayed in Figure 4. Module 1 consists of ten genes, including mitogen-activated protein kinase 1, peroxisome proliferator-activated receptor-gamma, snail family transcriptional repressor 1, interleukin-6, fibroblast growth factor 2, hepatocyte growth factor, estrogen receptor 1, mesenchymal-epithelial transition, hypoxia-inducible factor 1 subunit alpha, and SMAD2 with 42 interactions between them. KEGG pathway enrichment analysis of the genes in the module disclosed that nine out of ten genes excluding SANI1 were significantly enriched in cancer-related pathways. Furthermore, GO analysis revealed that all these genes are part of biological processes, such as positive regulation of transcription, and signal transduction pathway. Furthermore, Module 2 consists of 15 genes with 35 interactions between them. Among them, five out of 15 genes were involved in chromatid separation. Moreover, hub genes in the network were selected using network analyzer. Finally, we identified eight hub genes that pass the cutoff criteria: MAPK1, IL6, FGF2, SMAD2, SNAI1, DICER1, CDK6, and HGF.

Hub Gene Survival Analysis

For each hub gene, Kaplan–Meier survival curves for patients with lung adenocarcinoma and squamous cell carcinoma were plotted, and the log-rank test was used to assess statistical significance. NSCLC patients were split into two groups with high expression of each gene (above median) and low expression (below or equal to the median). The survival analysis identified seven of the ten genes which were statistically significantly different between patients

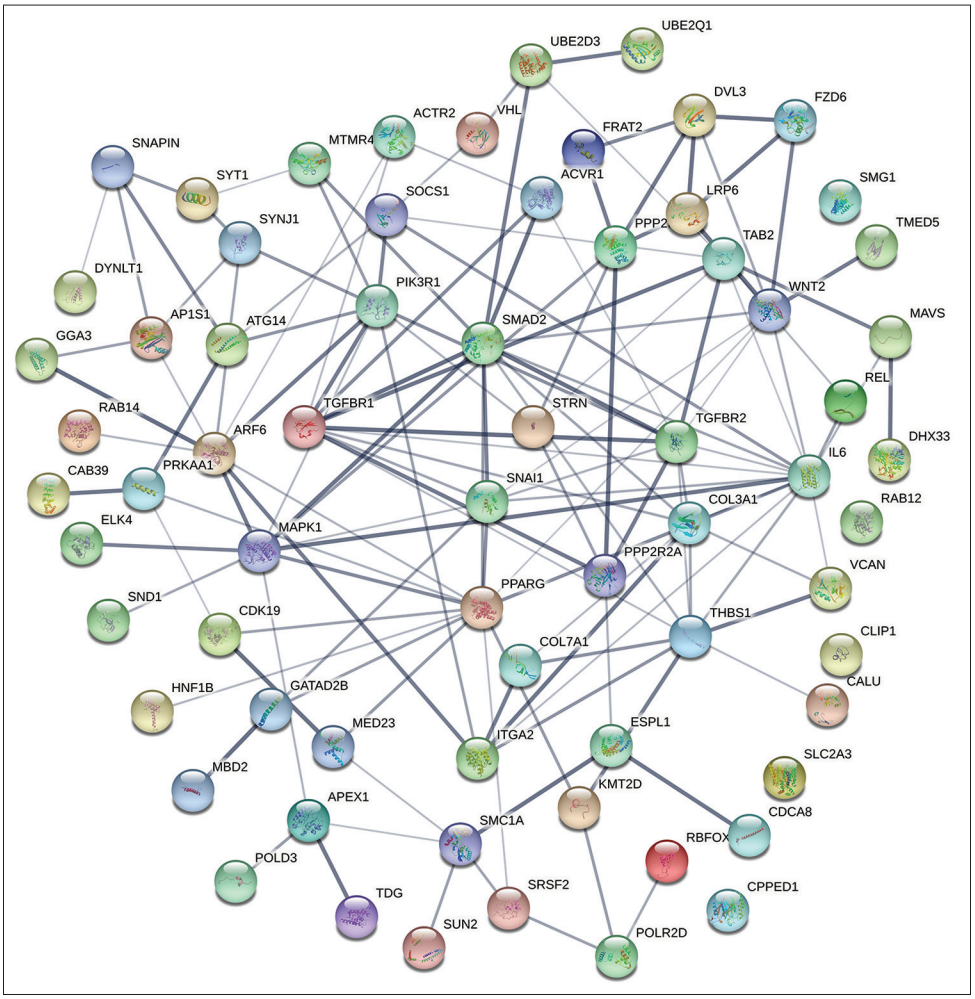

Figure 3: Protein-Protein Interaction [STRING-V11.5]

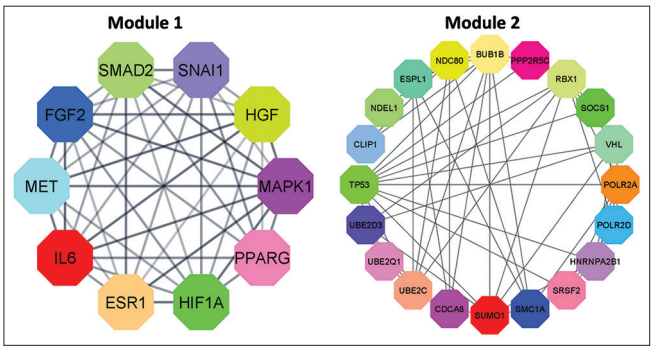

Figure 4: The highly interconnected genes in Module 1 and Module 2. Colours are arbitrary and each line indicates relationships between groups of genes. The first module consists of ten genes with 42 interactions, and the second module contains 20 genes with 35 interactions

with low and high expression. Consistent with these results, we also found that up-regulation of three genes (interleukin-6, snail family transcriptional repressor 1, and cell division protein kinase 6) was significantly associated with decreased survival and down-regulation of five genes (estrogen Receptor 1, fibroblast growth factor 2, SMAD family member 2, hepatocyte growth factor, DICER1, and hepatocyte growth factor) was significantly associated with poor prognosis [Figure 5]. Among the entire hub gene, SMAD2 showed the lowest HR (0.67) with a p-value of 2.9E-06. Taken together, this study suggests that these eight genes could hold promise as prognostic biomarkers of NSCLC. Among the eight most significant gene networks, the results indicated that IL6 genes tend to be higher degree nodes in the large PPI network.

In addition, we performed three genes set (IL6, SNAI1, and CDK6) signature analyses to find a panel of genes whose expression patterns correlated with overall survival. We found statistically significant overall survival between the high and low expressions group. Compared to each of the three genes alone, the HR for overall survival was higher for the combined three genes group with high prognostic power [Figure 6].

Identification of Potential miRNA Biomarkers for NSCLC

Kaplan–Meier survival analysis was used for the 12 differentially expressed miRNAs on 871 lung adenocarcinoma or lung squamous cell carcinoma patients. The results identified miRNAs (miR-140, miR-29c, and miR-199a) with p-values lower than 0.05 and HR 0.69 [Figure 7]. However, the other nine miRNAs did not have significant HRs. In addition, further analysis suggested that the up-regulation

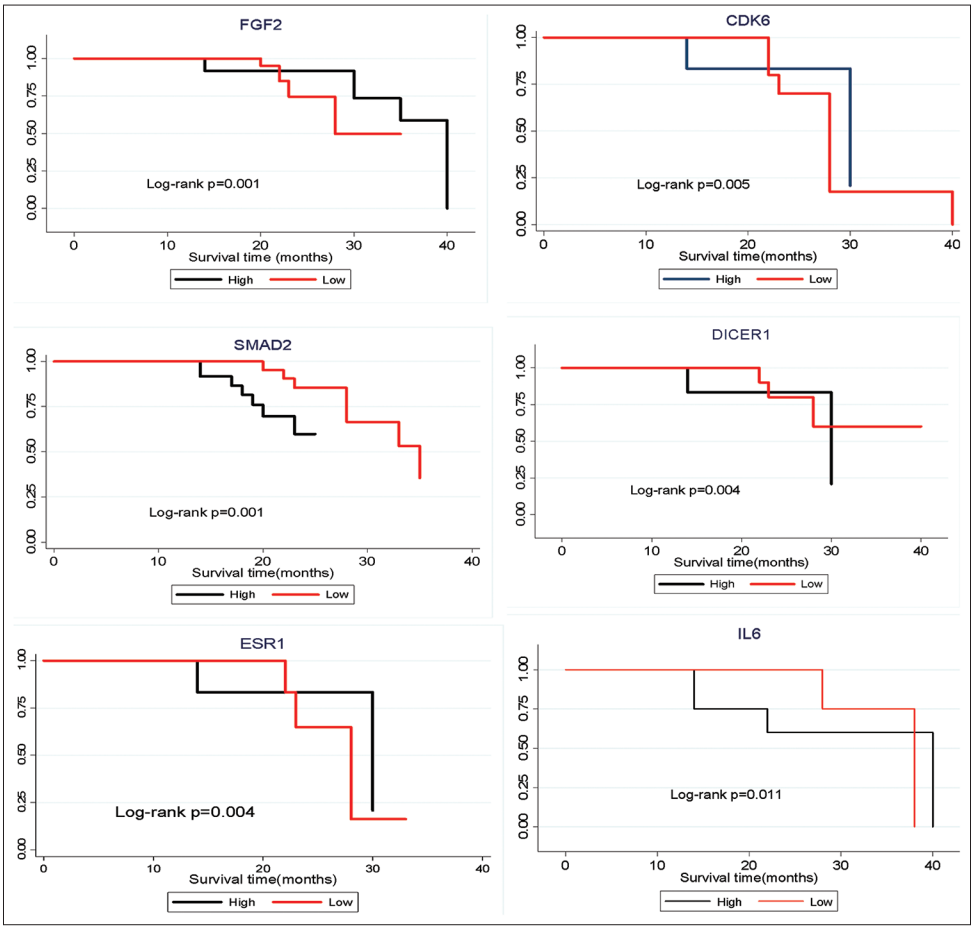

Figure 5: Kaplan–Meier survival plots for six genes. The black curve shows the survival of patients with up-regulated genes and the red curve shows the survival of those with down-regulated genes

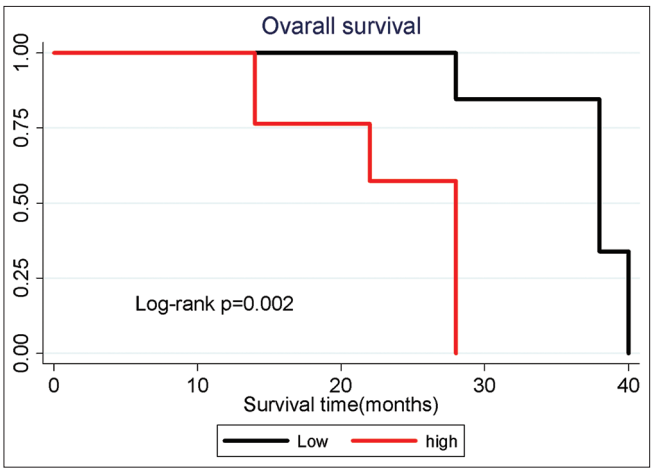

Figure 6: A combined survival curve analysis from the panel of three genes (IL6, SNAI1, and CDK6). Black colour indicates low expression and red colour shows high expression

of these three miRNAs in NSCLC was significantly associated with improved outcomes. In lung adenocarcinoma and lung squamous cell carcinoma tumour tissue samples, miR-140-3p expression was significantly down-regulated. The other miRNA that was found to be significantly down-regulated in LUSC patients was miR-29c-3p, while it was not significantly different for LUAD [Figure 8]. In lung adenocarcinoma and lung squamous cell carcinoma samples, miR-199a-5p expression was significantly up-regulated. The areas under the ROC curves (AUC) for each of the three miRNAs were generated. Among the three miRNAs, miRNA-140-3p showed the highest AUC (0.81), whereas miR-199a-5p and miR-29c-3p showed AUC values of 0.69 and 0.75, respectively [Figure 9].

Discussion

Despite significant improvements in NSCLC therapy, the prognosis for patients with NSCLC remains poor. This is mainly because of the lack of early and sensitive biomarkers. We demonstrate that the blood and tissue miRNAs and mRNAs and their profiles can be developed as biomarkers in the diagnosis and prognosis of NSCLC. This study examined four miRNA expression datasets and observed a substantial overlap of 12 differentially expressed miRNAs and 330 target genes were identified for 12 differentially expressed miRNAs.

In addition, functional enrichment analysis of the candidate miRNAs target genes disclosed some important biological processes and molecular functions. Dysregulated SMAD signaling has been linked to cancer.[35-37] Network analysis identified that SMAD3 appears in most enrichment pathways.

The main terms of molecular functions were related to transcription factor binding, transcription regulatory activity,

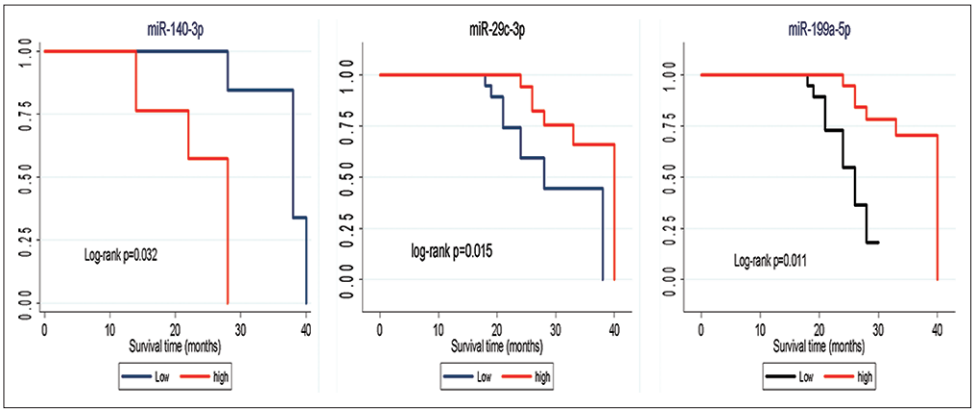

Figure 7: Kaplan–Meier survival curves for miRNAs associated with overall survival. The black line shows down-regulated miRNA in NSCLC tissue and the red line shows up-regulated miRNA in NSCLC tissue

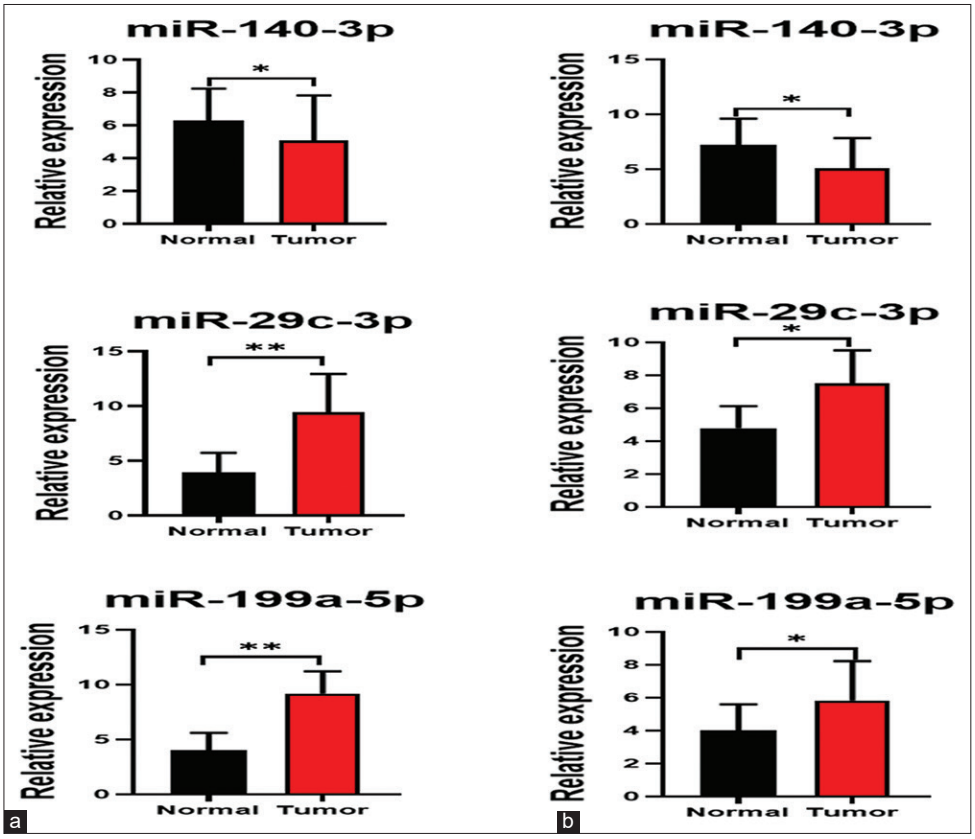

Figure 8: Relative expression of miRNAs. (a) LUAD versus control and (b) LUSC versus controls

protein kinase activity, and chromatin binding. In molecular function, 5′-deoxyribose-5-phosphate lyase activity was the most highly enriched. The enzyme 5′-deoxyribose-5-phosphate lyase is involved in base excision repair, which is a crucial step in the repair of DNA single-strand breaks and a major line of defence against cancer growth.[38,39] In addition, we observe a prominent enrichment for the RISC complex. MicroRNAs regulate gene expression by RNA-induced silencing complex (RISC) to a target mRNA, leading to translational repression and/or mRNA degradation. This indicated the possible role of epigenetic control of gene expression in NSCLC. We also identified eight genes (MPAL1, IL6, Fibroblast Growth Factor 2, mothers against decapentaplegic homolog 2, DICER1, cell division protein kinase 6, hepatocyte growth factor, and FGF2) that we considered to be major hubs. Among these genes, MAPK1 and IL6 are already known to be implicated in cancer. These eight hub genes are proposed as blood-born mRNA biomarkers for NSCLC and should be experimentally validated in the future work.

This study identified three miRNAs that could serve as biomarkers in the diagnosis of NSCLC: miR-140, miR-29c, and miR-199a. In addition, various studies showed that miR-140-3p functioned as a tumour suppressor in various malignancies, and miR-140-3p expression has been found to be down-regulated in

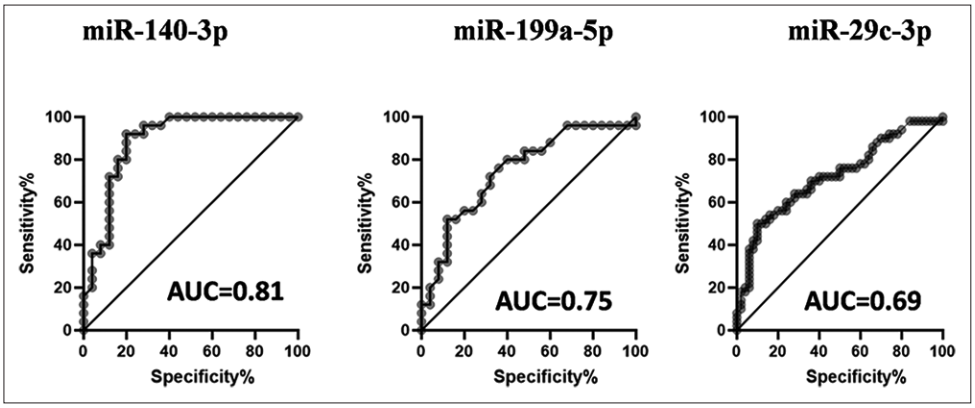

Figure 9: Receiver operating characteristic curve of three miRNAs. AUC, area under the ROC curve

a variety of cancers. This is in line with the results of the present study where decreasing levels of miR-140-3p were found in four separate NSCLC datasets. Recently, Huang *et al.* showed that the expression of miR-140-5p was decreased in the SCLC tissues, significantly associated with overall survival and cancer stage.[40] An additional study found that miR-140-3p was down-regulated in NSCLC tumours; compared with adjacent normal lung tissue.[41]

In addition, miR-29c-3p also showed the potential of using it as a non-invasive blood-based biomarker for the detection of NSCLC. Studies demonstrated that miR-29c-3p suppresses colon cancer cell growth and suppresses hepatocellular carcinoma tumour progression.[42,43] In laryngeal squamous cell carcinoma, low expression of miR-29c-3p is positively correlated with decreased survival.[44] Furthermore, decreased expression of miR-29c-3p suppresses endometrial cancer cells proliferation and tumour growth.[45]

Chen *et al.* identified the miR-199a-5p as a tumour suppressor in triple-negative breast cancer.[46] In addition, one study showed that miR-199a-5p targets MAP3K11 and suppresses NSCLC progression.[47] Zhu *et al.* demonstrated that miR-199a-5p suppresses colorectal tumour growth.[48] An additional study showed that miR-199a-5p targets SNAI1 and inhibits papillary thyroid carcinoma progression.[49]

One of the limitations of this study was that it only compared NSCLC patients with healthy controls, but not with other types of malignancy or lung diseases. This study identified mRNA and miRNA as biomarkers for NSCLC, however, not be able to differentiate between non-small-cell lung carcinoma and other carcinomas. The future research should be designed to include more samples to assess the prognostic and diagnostic potential of the identified miRNAs and mRNAs.

Conclusion

In this study, we identified 12 miRNAs that are down-regulated in the blood and tissue of NSCLC. Based on bioinformatics methods, three miRNAs (miR-140-3p, -29c, and -199a) are identified as potential biomarkers. Further studies with a larger number of samples are needed.

Acknowledgments

Not applicable.

Copyright and Permission Statement

We confirm that the materials included in this chapter do not violate copyright laws. All original sources have been appropriately acknowledged and/or referenced.

Funding

This study did not receive any specific research grant through a private, public, or non-profit funding agency.

References

1. GLOBOCAN 2020: New Global Cancer Data | UICC. Available: from: https://www.uicc.org/news/globocan-2020-new-global-cancer-data. [Last accessed on 2021 Oct 16].
2. Sánchez-Jiménez C, Ludeña MD, Izquierdo JM. T-cell intracellular antigens function as tumor suppressor genes. Cell Death Dis 2015;6:e1669.
3. Browse the Tables and Figures - SEER Cancer Statistics Review (CSR) 1975-2012. In: SEER. Available: from: https://www.seer.cancer.gov/archive/csr/1975_2012/browse_csr.php?sectionSEL=15&pageSEL=sect_15_table.14. [Last accessed on 2021 Sep 03].
4. Jacob F, Monod J. Genetic regulatory mechanisms in the synthesis of proteins. J Mol Biol 1961;3:318-56.
5. Lee RC, Feinbaum RL, Ambros V. The C. elegans heterochronic gene lin-4 encodes small RNAs with antisense complementarity to lin-14. Cell 1993;75:843-54.
6. Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R. Control of translation and mRNA degradation by miRNAs and siRNAs. Genes Dev 2006;20:515-24.
7. Catto JW, Alcaraz A, Bjartell AS, De Vere White R, Evans CP, Fussel S, *et al.* MicroRNA in prostate, bladder, and kidney cancer: A systematic review. Eur Urol 2011;59:671-81.
8. Iqbal MA, Arora S, Prakasam G, Calin GA, Syed MA. MicroRNA in lung cancer: Role, mechanisms, pathways and therapeutic relevance. Mol Aspects Med 2019;70:3-20.
9. Ma J, Lin Y, Zhan M, Mann DL, Stass SA, Jiang F. Differential miRNA expressions in peripheral blood mononuclear cells for diagnosis of lung cancer. Lab Invest 2015;95:1197-206.
10. Hennessey PT, Sanford T, Choudhary A, Mydlarz WW, Brown D, Adai AT, *et al.* Serum microRNA biomarkers for detection of non-small cell lung cancer. PLoS One 2012;7:e32307.
11. Shen J, Todd NW, Zhang H, Yu L, Lingxiao X, Mei Y, *et al.* Plasma microRNAs as potential biomarkers for non-small-cell lung cancer. Lab Invest 2011;91:579-87.
12. Zhao H, Shen J, Medico L, Wang D, Ambrosone CB, Liu S. A pilot study

of circulating miRNAs as potential biomarkers of early stage breast cancer. PLoS One 2010;5:e13735.
13. Szafranska AE, Davison TS, John J, Cannon T, Sipos B, Maghnouj A, *et al.* MicroRNA expression alterations are linked to tumorigenesis and non-neoplastic processes in pancreatic ductal adenocarcinoma. Oncogene 2007;26:4442-52.
14. Galka-Marciniak P, Urbanek-Trzeciak MO, Nawrocka PM, Dutkiewicz A, Giefing M, Lewandowska MA, *et al.* Somatic mutations in miRNA genes in lung cancer-potential functional consequences of non-coding sequence variants. Cancers (Basel) 2019;11:793.
15. Xue WX, Zhang MY, Rui Li, Liu X, Yin YH, Qu YQ. Serum miR-1228-3p and miR-181a-5p as noninvasive biomarkers for non-small cell lung cancer diagnosis and prognosis. Biomed Res Int 2020;2020:9601876.
16. Riveiro ME, Astorgues-Xerri L, Vazquez R, Frapolli R, Kwee I, Rinaldi A, *et al.* OTX015 (MK-8628), a novel BET inhibitor, exhibits antitumor activity in non-small cell and small cell lung cancer models harboring different oncogenic mutations. Oncotarget 2016;7:84675-87.
17. Henderson LJ, Coe BP, Lee EH, Girard L, Gazdar AF, Minna JD, *et al.* Genomic and gene expression profiling of minute alterations of chromosome arm 1p in small-cell lung carcinoma cells. Br J Cancer 2005;92:1553-60.
18. Li G, Gao Y, Cui Y, Zhang T, Cui R, Jiang Y, *et al.* Overexpression of CD44 is associated with the occurrence and migration of non-small cell lung cancer. Mol Med Rep 2016;14:3159-67.
19. Morris S, Vachani A, Pass HI, Rom WN, Ryden K, Weiss GJ, *et al.* Whole blood FPR1 mRNA expression predicts both non-small cell and small cell lung cancer. Int J Cancer 2018;142:2355-62.
20. Hosseini SM, Soltani BM, Tavallaei M, Mowla SJ, Tafsiri E, Bagheri A, *et al.* Clinically significant dysregulation of hsa-miR-30d-5p and hsa-let-7b expression in patients with surgically resected non-small cell lung cancer. Avicenna J Med Biotechnol 2018;10:98-104.
21. Yang F, Wei K, Qin Z, Liu W, Shao C, Wang C, *et al.* MiR-598 suppresses invasion and migration by negative regulation of derlin-1 and epithelial-mesenchymal transition in non-small cell lung cancer. Cell Physiol Biochem 2018;47:245-56.
22. Olivieri F, Capri M, Bonafè M, Morsiani C, Jung HJ, Spazzafumo L, *et al.* Circulating miRNAs and miRNA shuttles as biomarkers: Perspective trajectories of healthy and unhealthy aging. Mech Ageing Dev 2017;165:162-70.
23. Huang R, Gao L. Identification of potential diagnostic and prognostic biomarkers in non-small cell lung cancer based on microarray data. Oncol Lett 2018;15:6436-42.
24. Edgar R, Domrachev M, Lash AE. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. Nucleic Acids Res 2002;30:207-10.
25. Barrett T, Wilhite SE, Ledoux P, Evangelista C, Kim IF, Tomashevsky M, *et al.* NCBI GEO: Archive for functional genomics data sets – Update. Nucleic Acids Res 2013;41:D991-5.
26. Asakura K, Kadota T, Matsuzaki J, Yoshida Y, Yamamoto Y, Nakagawa K, *et al.* A miRNA-based diagnostic model predicts resectable lung cancer in humans with high accuracy. Commun Biol 2020;3:134.
27. Qu L, Li L, Zheng X, Fu H, Tang C, Qin H, *et al.* Circulating plasma microRNAs as potential markers to identify EGFR mutation status and to monitor epidermal growth factor receptor-tyrosine kinase inhibitor treatment in patients with advanced non-small cell lung cancer. Oncotarget 2017;8:45807-24.
28. Li LL, Qu LL, Fu HJ, Zheng XF, Tang CH, Li XY, *et al.* Circulating microRNAs as novel biomarkers of ALK-positive nonsmall cell lung cancer and predictors of response to crizotinib therapy. Oncotarget 2017;8:45399-414.
29. Pu HY, Xu R, Zhang MY, Yuan LJ, Hu JY, Huang GL, *et al.* Identification of microRNA-615-3p as a novel tumor suppressor in non-small cell lung cancer. Oncol Lett 2017;13:2403-10.
30. Dweep H, Sticht C, Pandey P, Gretz N. miRWalk - Database: Prediction of possible miRNA binding sites by "walking" the genes of three genomes. J Biomed Inform 2011;44:839-47.
31. Mi H, Huang X, Muruganujan A, Tang H, Mills C, Kang D, *et al.* PANTHER version 11: Expanded annotation data from Gene Ontology and Reactome pathways, and data analysis tool enhancements. Nucleic Acids Res 2017;45:D183-9.
32. von Mering C, Huynen M, Jaeggi D, Schmidt S, Bork P, Snel B. STRING: A database of predicted functional associations between proteins. Nucleic Acids Res 2003;31:258-61.
33. Nagy Á, Lánczky A, Menyhárt O, Győrffy B. Validation of miRNA prognostic power in hepatocellular carcinoma using expression data of independent datasets. Sci Rep 2018;8:9227.
34. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: A web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017;45:W98-102.
35. Bello-DeOcampo D, Tindall DJ. TGF-betal/Smad signaling in prostate cancer. Curr Drug Targets 2003;4:197-207.
36. Xie W, Rimm DL, Lin Y, Shih WJ, Reiss M. Loss of Smad signaling in human colorectal cancer is associated with advanced disease and poor prognosis. Cancer J 2003;9:302-12.
37. Cui Y, Song Y, Yan S, Cao M, Huang J, Jia D, *et al.* CUEDC1 inhibits epithelial-mesenchymal transition via the TβRI/Smad signaling pathway and suppresses tumor progression in non-small cell lung cancer. Aging (Albany NY) 2020;12:20047-68.
38. García-Díaz M, Bebenek K, Kunkel TA, Blanco L. Identification of an intrinsic 5′-deoxyribose-5-phosphate lyase activity in human DNA polymerase lambda: A possible role in base excision repair. J Biol Chem 2001;276:34659-63.
39. Longley MJ, Prasad R, Srivastava DK, Wilson SH, Copeland WC. Identification of 5′-deoxyribose phosphate lyase activity in human DNA polymerase gamma and its role in mitochondrial base excision repair *in vitro*. Proc Natl Acad Sci U S A 1998;95:12244-8.
40. Huang H, Wang Y, Li Q, Fei X, Ma H, Hu R. miR-140-3p functions as a tumor suppressor in squamous cell lung cancer by regulating BRD9. Cancer Lett 2019;446:81-9.
41. Kong XM, Zhang GH, Huo YK, Zhao XH, Cao DW, Guo SF, *et al.* MicroRNA-140-3p inhibits proliferation, migration and invasion of lung cancer cells by targeting ATP6AP2. Int J Clin Exp Pathol 2015;8:12845-52.
42. Chen G, Zhou T, Li Y, Yu Z, Sun L. p53 target miR-29c-3p suppresses colon cancer cell invasion and migration through inhibition of PHLDB2. Biochem Biophys Res Commun 2017;487:90-5.
43. Wu H, Zhang W, Wu Z, Liu Y, Shi Y, Gong J, *et al.* miR-29c-3p regulates DNMT3B and LATS1 methylation to inhibit tumor progression in hepatocellular carcinoma. Cell Death Dis 2019;10:48.
44. Fang R, Huang Y, Xie J, Zhang J, Ji X. Downregulation of miR-29c-3p is associated with a poor prognosis in patients with laryngeal squamous cell carcinoma. Diagn Pathol 2019;14:109.
45. Van Sinderen M, Griffiths M, Menkhorst E, Niven K, Dimitriadis E. Restoration of microRNA-29c in type I endometrioid cancer reduced endometrial cancer cell growth. Oncol Lett 2019;18:2684-93.
46. Chen J, Shin VY, Siu MT, Ho JC, Cheuk I, Kwong A. miR-199a-5p confers tumor-suppressive role in triple-negative breast cancer. BMC Cancer 2016;16:887.
47. Li Y, Wang D, Li X, Shao Y, He Y, Yu H, *et al.* MiR-199a-5p suppresses non-small cell lung cancer via targeting MAP3K11. J Cancer 2019;10:2472-9.
48. Zhu QD, Zhou QQ, Dong L, Huang Z, Wu F, Deng X. MiR-199a-5p inhibits the growth and metastasis of colorectal cancer cells by targeting ROCK1. Technol Cancer Res Treat 2018;17:1533034618775509.
49. Ma S, Jia W, Ni S. miR-199a-5p inhibits the progression of papillary thyroid carcinoma by targeting SNAI1. Biochem Biophys Res Commun 2018;497:181-6.