Journal of Advanced Biomedical Sciences
https://jabs.fums.ac.ir/
Online ISSN: 2783-1523
Fasa University of Medical Sciences

Cigarette Smoke-Induced Transcriptomic Alterations and Angiogenesis in Non-Small Cell Lung Cancer: An Integrative Analysis

Kianoush Mohammadi, Reza Safaralizadeh

Department of Animal Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran

Article Info

Article Type:
Original Article

Article history:
Received
18 Mar 2025
Received in revised form
27 Apr 2025
Accepted
06 May 2025
Published online
10 Jun 2025

Publisher
Fasa University of Medical Sciences

Abstract

Background & Objectives: Cigarette smoke is a major risk factor for non-small cell lung cancer (NSCLC) and plays a pivotal role in tumor initiation and angiogenesis. This study aimed to elucidate the molecular mechanisms through which cigarette smoke influences angiogenesis in NSCLC by integrating transcriptomic data, with a particular emphasis on the regulatory role of microRNA-1 (miR-1) and its downstream targets.

Materials & Methods: We analyzed the Gene Expression Omnibus (GEO) dataset GSE290190, comprising gene expression profiles from 18 samples with different smoking statuses (9 normal and 9 tumor tissues). Differential expression analysis, Gene Set Enrichment Analysis (GSEA), Gene Ontology (GO) enrichment, and Protein–Protein Interaction (PPI) network analysis were conducted to identify critical genes and signaling pathways. Statistical analyses were employed to determine differentially expressed genes (DEGs) and to assess their biological relevance.

Results: Differential expression analysis identified 2,449 DEGs between normal and tumor tissues, with significant enrichment in angiogenesis, cell cycle regulation, and DNA repair pathways. Key angiogenesis-related genes—VEGFC, FGF2, and ANGPT1—were recognized as direct targets of miR-1. GSEA and GO analyses revealed marked alterations in biological processes such as chromosome segregation, mitotic nuclear division, and extracellular matrix organization. PPI network analysis identified E2F7, PLK1, and TOP2A as hub genes, suggesting their potential roles as key regulators in cell cycle progression and tumorigenesis.

Conclusion: This study highlights the transcriptomic heterogeneity of NSCLC and proposes miR-1 and its downstream targets—VEGFC, FGF2, and ANGPT1—as promising biomarkers and therapeutic targets. However, further validation using larger datasets and functional assays is essential to confirm these findings and facilitate their clinical translation.

Keywords: Non-Small Cell Lung Cancer; Cigarette Smoke; Angiogenesis; MicroRNA-1; Transcriptomic Analysis

Cite this article: Mohammadi K, Safaralizadeh R. Cigarette Smoke-Induced Transcriptomic Alterations and Angiogenesis in Non-Small Cell Lung Cancer: An Integrative Analysis. J Adv Biomed Sci. 2025; 15(3): 260-273.
DOI: 10.18502/jabs.v15i3.18951

Introduction

Bronchogenic carcinoma, commonly known as lung cancer, refers to malignancies originating in the lung parenchyma or bronchi. It is among the leading causes of cancer-related mortality worldwide (1). Paradoxically, lung cancer was relatively uncommon in the early 20th century; however, its incidence has surged dramatically in recent decades, primarily due to increased cigarette consumption. It is estimated

Corresponding Author: Reza Safaralizadeh, Department of Animal Biology, Faculty of Natural Sciences, University of Tabriz, Tabriz, Iran
Email: safaralizadeh@tabrizu.ac.ir

CC BY NC

that smoking accounts for approximately 90% of lung cancer cases, while exposure to other carcinogens, such as asbestos, further elevates the risk (2,3). Approximately 80% of lung cancers are classified as non-small cell lung cancer (NSCLC), which generally exhibits slower growth compared to other subtypes. Nevertheless, this form of cancer may metastasize to distant organs, including the brain and bones, before clinical symptoms become evident (4).

Lung cancer staging is most commonly conducted using the TNM classification system, which evaluates tumor size and local invasion (T), lymph node involvement (N), and distant metastasis (M). The clinical progression of the disease is typically categorized as follows:

1. **Localized (Stage I–II):** The tumor is confined to the lung, with no involvement of lymph nodes or distant organs.
2. **Regional (Stage III):** The malignancy has spread to nearby lymph nodes or adjacent structures but remains within the thoracic cavity.
3. **Distant (Stage IV):** Metastases have occurred in remote organs such as the brain, bones, or liver (5).

Angiogenesis—the formation of new blood vessels from pre-existing vasculature—is a critical process driven by the metabolic demands of both normal and neoplastic tissues. In the context of lung cancer, angiogenesis facilitates tumor growth and invasion and is closely associated with metastasis and immune modulation within the tumor microenvironment (6). Cigarette smoke has been shown to promote tumorigenesis by enhancing cellular proliferation, migration, invasion, and angiogenesis (7). Vascular endothelial growth factor (VEGF), a key mediator of tumor angiogenesis and progression, is upregulated in part by the downregulation of microRNA-1 (miR-1) in pulmonary endothelial cells. MiR-1 has been identified as a prognostic marker for NSCLC and may inform therapeutic monitoring and patient management strategies (8). Furthermore, cigarette smoke has been reported to activate the angiogenic switch in lung cancer by suppressing miR-1 expression, thereby positioning miR-1 as a potential biomarker for disease initiation and recurrence (9).

Overexpression of miR-1 has also been associated with increased sensitivity to cisplatin (DDP) through autophagy-related 3 (ATG3)-mediated autophagy, thereby enhancing the efficacy of chemotherapeutic agents and highlighting its potential as a therapeutic target (10). The selection of miR-1 as the focal point of this study was based on evidence from the referenced dataset, which identified cigarette smoke-induced degradation of mature miR-1 as a central regulatory mechanism.

Despite accumulating evidence implicating cigarette smoke in NSCLC pathogenesis, the precise molecular mechanisms underlying its angiogenic effects—particularly those mediated by miRNAs—remain inadequately characterized. There is a notable paucity of integrative transcriptomic studies specifically examining the interplay between smoking status and miR-1 regulation, underscoring a critical gap in current knowledge. Elucidating these molecular pathways is imperative for the identification of early diagnostic biomarkers and the development of personalized therapeutic strategies (11).

Materials and Methods

This study employed a multi-step methodology encompassing data preprocessing, statistical analysis, and visualization. We analyzed gene expression alterations in 26,485 genes across 18 samples derived from the publicly accessible GEO dataset GSE290190, available at: https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE290190. This dataset specifically investigates the dysregulation of endothelial microRNA-1 (miR-1) in NSCLC patients as a consequence of cigarette smoke exposure. The characteristics of the analyzed samples, including patient smoking status, are summarized in Table 1 below.

Table 1. Characteristics of Samples and Patients

Sample ID	Patient ID	Tissue Type	Smoking Status	Sample Type
GSM8808012	387	Normal Tissue	Former Smoker	NF
GSM8808013	387	Tumor Tissue	Former Smoker	TF
GSM8808014	395	Normal Tissue	Former Smoker	NF
GSM8808015	395	Tumor Tissue	Former Smoker	TF
GSM8808016	397	Normal Tissue	Current Smoker	NC
GSM8808017	397	Tumor Tissue	Current Smoker	TC
GSM8808018	406	Normal Tissue	Current Smoker	NC
GSM8808019	406	Tumor Tissue	Current Smoker	TC
GSM8808020	411	Normal Tissue	Never Smoker	NN
GSM8808021	411	Tumor Tissue	Never Smoker	TN
GSM8808022	417	Normal Tissue	Former Smoker	NF
GSM8808023	417	Tumor Tissue	Former Smoker	TF
GSM8808024	452	Normal Tissue	Current Smoker	NC
GSM8808025	452	Tumor Tissue	Current Smoker	TC
GSM8808026	529	Normal Tissue	Never Smoker	NN
GSM8808027	529	Tumor Tissue	Never Smoker	TN
GSM8808028	532	Normal Tissue	Never Smoker	NN
GSM8808029	532	Tumor Tissue	Never Smoker	TN

All necessary preprocessing and analytical procedures were conducted using R software (version 4.2.2). The required packages were first installed and subsequently loaded. The following packages were utilized: “DESeq2” (v1.46.0), “ggplot2” (v3.5.1), “EnhancedVolcano” (v1.24.0), “pheatmap” (v1.0.12), “dplyr” (v1.1.4), “tidyverse” (v2.0.0), “org.Hs.eg.db” (v3.20.0), and “STRINGdb” (v2.18.0). These packages were installed and loaded via the BiocManager package (v1.30.25), which facilitates the management of Bioconductor libraries (12–19).

The dataset was then imported into R, and analyses were carried out at four hierarchical levels, encompassing four distinct sample types. Initially, to compare tumor and normal groups, samples were categorized into two primary groups: Normal (N) and Tumor (T). To assess the potential influence of smoking status on the observed outcomes, the sample groups were further stratified into three subgroups: Normal Current Smoker vs. Tumor Current Smoker (NC vs. TC), Normal Former Smoker vs. Tumor Former Smoker (NF vs. TF), and Normal Never Smoker vs. Tumor Never Smoker (NN vs. TN).

Differential Expression Analysis

Differential expression analysis was performed by constructing a DESeqDataSet object, followed by normalization using the median-of-ratios method. This normalization approach ensures that observed variations in gene expression are attributable to underlying biological factors rather than technical artifacts. A properly structured conditions vector was created, and differential expression was assessed using the DESeq2 package, enabling systematic comparisons across sample groups (20).

Following this, a filtering step was applied to retain only statistically significant results, specifically genes with an adjusted p-value (padj) <0.05. The resulting output included critical metrics such as baseMean, log2FoldChange, lfcSE, stat, p-value, and padj, each providing insight into expression magnitude and statistical significance.

To visualize differentially expressed genes, two graphical representations were employed: the volcano plot and the heatmap. The volcano plot incorporated all genes—regardless of significance level—and displayed

log2FoldChange on the x-axis and –log10(p-value) on the y-axis. A p-value cutoff of 0.05 and a fold-change threshold of 1 were applied (pCutoff = 0.05; FCcutoff = 1) (21). For the heatmap, only significantly differentially expressed genes (padj < 0.05) were considered. The dataset was refined to include the top 50 most significantly altered genes (i.e., those with the lowest padj values) to enhance interpretability and minimize overplotting (top_genes <- rownames(res[order(res$padj),]) [1:50]).

Additionally, Principal Component Analysis (PCA) was conducted using variance-stabilized transformed data (vst(dds)) to assess global expression patterns across samples. This analysis was stratified by the "condition" variable, allowing for clear visualization of sample clustering. The percentage of variance explained by each principal component (percentVar) was computed to elucidate the primary axes of variation in the dataset. This technique provides insight into the overarching gene expression structure under different experimental conditions (22).

Gene Set Enrichment Analysis (GSEA)

To elucidate the biological significance of the differentially expressed genes (DEGs), Gene Set Enrichment Analysis (GSEA) was performed using the clusterProfiler package (v4.14.4) (23). The gseGO function was employed to map Ensembl gene identifiers to Gene Ontology (GO) terms, utilizing the org.Hs.eg.db annotation database for *Homo sapiens* (24). The minimum gene set size was set to 1, and a p-value threshold of 0.05 was used to identify significantly enriched terms. A dot plot was generated to visually summarize enriched GO categories. Furthermore, gene name-to-Ensembl ID conversion was conducted using the biomaRt package (v2.62.1).

Functional Annotation and Pathway Analysis

To investigate the functional roles of DEGs, we initiated the analysis by loading the org.Hs.eg.db package (v3.20.0). Enrichment analysis was then carried out using the enrichGO function, with p-value and q-value thresholds both set at 0.05. The corresponding enrichment results were visualized using the dotplot function (25). Gene names served as the input for this analysis, with Ensembl IDs converted via the biomaRt package.

In parallel, KEGG pathway enrichment analysis was conducted using the enrichKEGG function, specifying *Homo sapiens* (organism code: "hsa"). This analysis incorporated significantly expressed genes (padj < 0.05) and applied consistent p-value and q-value cutoffs of 0.05 (26). Resulting pathways were similarly visualized using dot plots generated via the clusterProfiler package. The conversion from gene symbols to Entrez IDs was accomplished using the mapIds function in conjunction with the org.Hs.eg.db database. Collectively, this rigorous multi-tiered approach enabled the elucidation of key biological processes and pathways associated with the observed transcriptional changes.

Protein-Protein Interaction (PPI) Network Analysis

To identify functional interactions among proteins encoded by DEGs and to detect potential hub genes, a Protein–Protein Interaction (PPI) network analysis was undertaken using the STRINGdb package (v2.18.0), configured for *Homo sapiens* (NCBI taxonomy ID: 9606). This package was installed and managed through BiocManager. A STRINGdb object was instantiated with a confidence score threshold of 400, thereby focusing on high-confidence interactions.

The analysis utilized Entrez IDs corresponding to the top 100 most significant genes, ranked by lowest adjusted p-values. Genes lacking interactions or displaying minimal connectivity were excluded from further visualization. The final PPI network was subsequently visualized and archived, providing a comprehensive overview of interaction dynamics among key DEGs (27).

Table 2. Filtering results for each comparison's 10 most important genes and their numerical values

Sample Type	Genes with Highest log2FoldChange	Genes with Lowest log2FoldChange	Genes with Lowest padj
T vs N	CST4 (7.79), LINC01249 (7.53), KRT75 (7.14), PNPLA5 (7.11), CST1 (7.02), LOC101927136 (6.77), SPINK1 (6.42), SPRR3 (6.11), ACTL8 (6.10)	HAS1 (-6.33), MIR663A (-5.92), SLC6A4 (-5.23), CSF3 (-5.19), IL13 (-5.14), LY6G6E (-5.09), LOC101927410 (-4.99), LINC01069 (-4.98), MT1A (-4.85)	ANKRD1 (3.36E-13), COL10A1 (4.37E-11), BMPER (1.45E-10), FCN3 (2.24E-10), RTKN2 (2.24E-10), CD36 (8.12E-10), SLITRK2 (1.54E-09), CST1 (1.82E-09), CST4 (1.28E-08)
NF vs TF	FAM138A (22.46), FAM138C (22.46), FAM138F (22.46), LINC00901 (21.03), DGKK (7.31), MOGAT1 (7.27), KCNJ3 (6.95), SLC6A4 (6.57), HAS1 (6.36)	RBM46 (-24.67), HOXC10 (-24.60), LINC01249 (-24.14), HOXD11 (-23.86), GABRG2 (-23.38), PRDM9 (-23.14), SALL1 (-22.78), MAGEB10 (-22.66), HOXA13 (-22.58)	FAM138A (1.42894E-25), FAM138C (1.42894E-25), FAM138F (1.42894E-25), RBM46 (4.88598E-22), SALL1 (2.28492E-19), MAGEB10 (5.56366E-14), LINC01193 (5.90815E-12), HOXC10 (8.04933E-12), SLITRK1 (3.05419E-11)
NC vs TC	PRDM9 (19.40), TFF1 (19.28), MT1A (9.58), ZFP42 (7.08), PLA2G2A (6.79), GRID2 (6.61), DEFA1B (6.14), DEFA1 (6.13), CD300LG (6.12)	CST4 (-7.50), SOX2 (-6.55), ANKRD34B (-6.20), SLC38A11 (-6.17), CNTNAP2 (-6.16), CST1 (-6.16), FOXE1 (-6.09), UGT1A7 (-6.01), UGT1A9 (-5.91)	SOX7 (9.14E-05), PDK4 (9.14E-05), CA4 (2.06E-04), CD36 (4.91E-04), AKAP2 (5.26E-04), LOC101928370 (5.26E-04), MT1A (5.26E-04), PALM2-AKAP2 (5.62E-04), S1PR1 (6.77E-04)
NN vs TN	LINC00383 (18.38), PSG4 (18.10), LOC101928272 (16.48), KLK8 (16.08), PRR9 (15.17), HAS1 (6.88), SEC14L3 (6.44), C11orf88 (6.34), HSD17B13 (4.67)	SALL1 (-23.21), OPRD1 (-22.60), HOXB13 (-21.23), LINC01249 (-20.35), MAGEA9B (-16.70), MAGEA9 (-16.70), CST4 (-8.32), RTBDN (-8.17), KIF1A (-7.68)	SALL1 (1.63E-19), LOC101928272 (1.52E-04), KIF1A (2.27E-04), MMP11 (2.27E-04), ABCA12 (1.46E-03), DNAJC22 (2.43E-03), TMEM63C (2.43E-03), RNF144A-AS1 (2.43E-03), COL22A1 (2.86E-03)

Results

In the comparison between normal (N) and tumor (T) samples (N vs. T), 2,449 DEGs were detected, of which 1,407 were up-regulated and 1,044 were down-regulated. The NF vs. TF contrast yielded 697 DEGs—314 up-regulated and 383 down-regulated—whereas the NC vs. TC analysis revealed 172 DEGs, 97 up-regulated and 75 down-regulated. Finally, 94 DEGs (25 up-regulated and 69 down-regulated) were identified in the NN vs. TN comparison (Table 2; Figures 1).

GSEA demonstrated significant enrichment of multiple biological processes between normal and tumor samples, with the most pronounced changes (expressed as $\log_2$-fold change) depicted in Figure 2. A comparable enrichment pattern was observed for the NF vs. TF comparison.

By contrast, the NC vs. TC analysis highlighted processes related to DNA-templated DNA replication, mitotic sister-chromatid segregation, cell-cycle checkpoint signalling, keratinocyte differentiation, mitotic nuclear division, nuclear chromosome segregation, regulation of chromosome segregation, and skin development. In the NN vs. TN comparison, processes associated with axoneme assembly, cilium-dependent motility, microtubule-bundle formation, and multiple chromosome-segregation terms were significantly enriched (Figure 2).

GO enrichment analysis of the N vs. T dataset returned 718 significantly enriched GO terms, many of which relate to mitotic events. Notably, "mitotic nuclear division" (GO:0140014), "chromosome segregation" (GO:0007059), "DNA replication" (GO:0006260; GO:0006261),

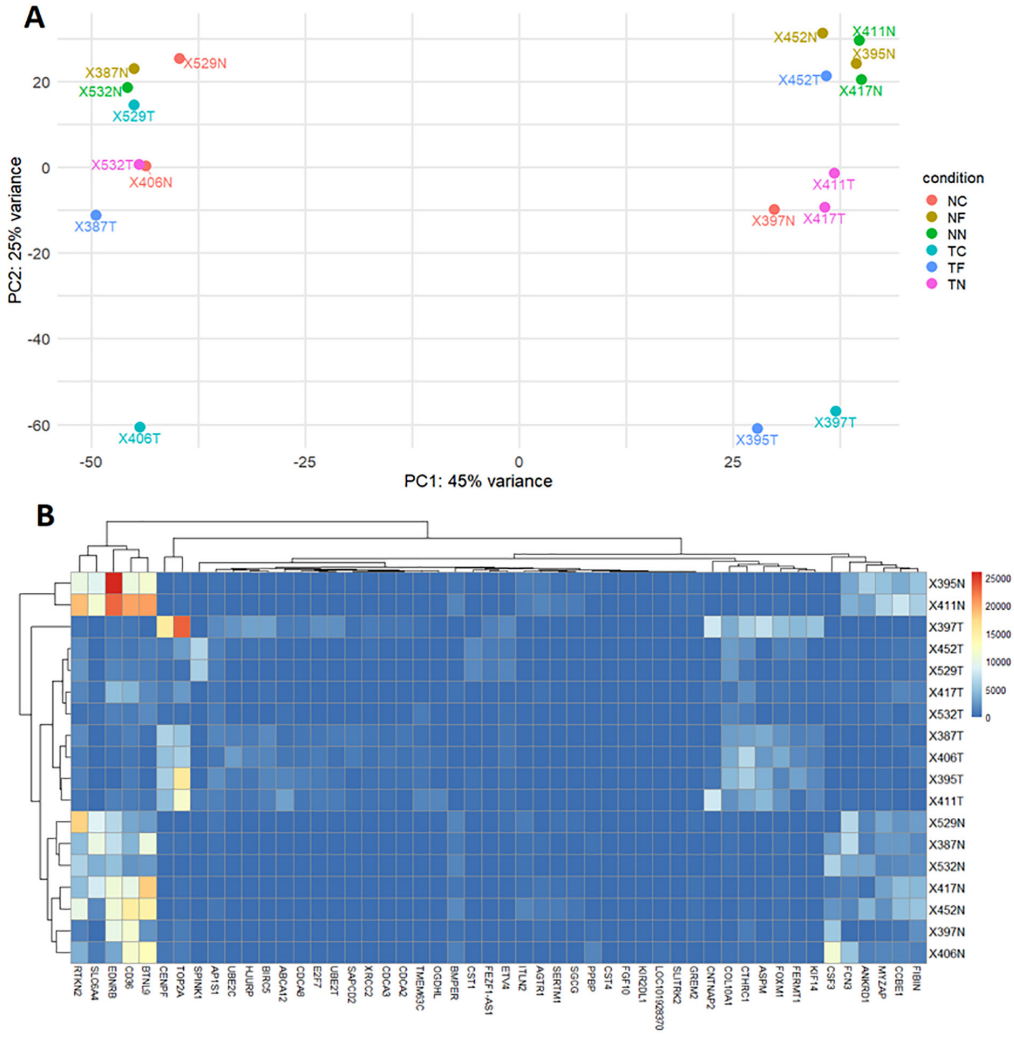


Figure 1. A: Principal Component Analysis (PCA) of Gene Expression Profiles. **B:** Heatmap of Differentially Expressed Genes (DEGs). Rows: Samples (N or T). Columns: Genes. Color key: Red = upregulated expression; Blue = downregulated expression

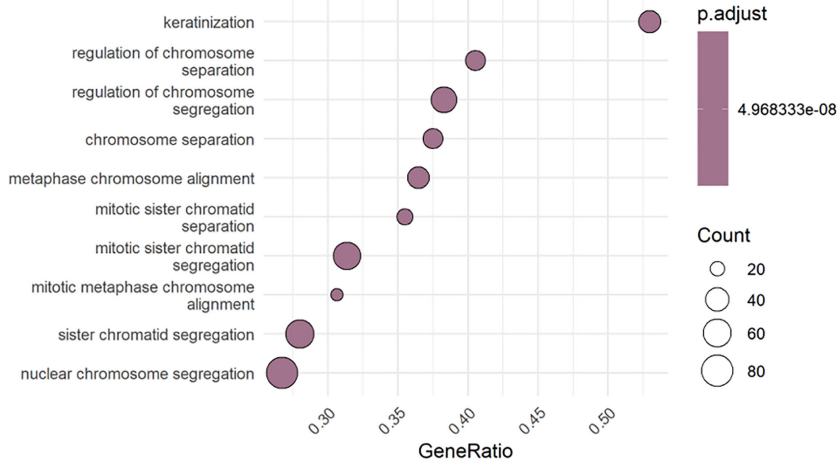


Figure 2. Gene Set Enrichment Analysis (GSEA) Normal vs. Tumor Samples (N vs. T)

and "regulation of cell cycle" (GO:0045786; GO:1901990) exhibited both low adjusted p-values and high gene ratios, underscoring the pivotal contribution of DEGs to cell-division regulation (Figure 3).

For the NF vs. TF contrast, GO terms describing chromosome segregation, homophilic cell adhesion, and the orchestration of nuclear and mitotic divisions were over-represented. The NC vs. TC comparison, however, was enriched for processes involving muscle-system development, axonogenesis, epithelial-cell proliferation, renal morphogenesis, ion-transport regulation, and amoeboid-type cell migration. In the NN vs. TN dataset, extracellular-matrix organisation and skeletal morphogenesis predominated, together with pathways that positively regulate angiogenesis and osteoblast proliferation. KEGG pathway enrichment analysis corroborated these findings, revealing robust pathway activation in the N vs. T comparison (Figure 4). Cell-cycle pathways predominated in the NF vs. TF group; the NC vs. TC contrast featured signalling modules such as cGMP–PKG, adrenergic signalling in cardiomyocytes, fatty-acid β-oxidation, vascular smooth-muscle contraction, salivary secretion, pyruvate metabolism, regulation of lipolysis in adipocytes, dilated cardiomyopathy, and fatty-acid metabolism; and the NN vs. TN analysis isolated pathways related to protein digestion and absorption (Figure 4).

To investigate the relationship among DEGs and to identify hub genes, we employed the STRING online database to construct the protein–protein interaction (PPI) network. In this network, each node represents a gene, whereas edges indicate predicted or known interactions (Figure 5).

Several key genes exhibiting significant interaction patterns in the comparison between normal (N) and tumor (T) samples were identified, including *E2F Transcription Factor 7* (E2F7), *Ubiquitin-Conjugating Enzyme E2 T* (UBE2T), *Kinesin Family Member 14* (KIF14), *Centromere Protein F* (CENPF), *Polo-Like Kinase 1* (PLK1), *Forkhead Box M1* (FOXM1), *Topoisomerase (DNA) II Alpha* (TOP2A), *Ubiquitin-Conjugating Enzyme E2 C* (UBE2C), *Cell Division Cycle Associated 8* (CDCA8), *Baculoviral IAP Repeat Containing 5* (BIRC5), *Cell Division Cycle Associated 2* (CDCA2), *Cell Division Cycle Associated 3* (CDCA3), *Abnormal Spindle Microtubule Assembly* (ASPM), and *Holliday Junction Recognition Protein* (HJURP).

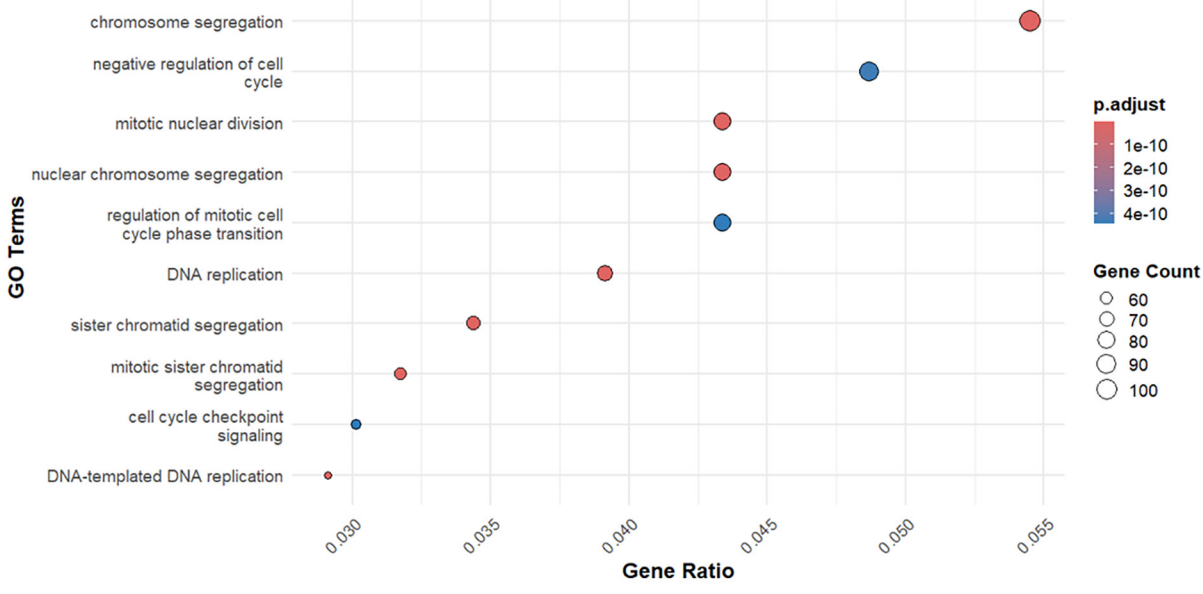


Figure 3. Gene Ontology (GO) Enrichment Analysis for Normal vs. Tumor Samples (N vs. T)

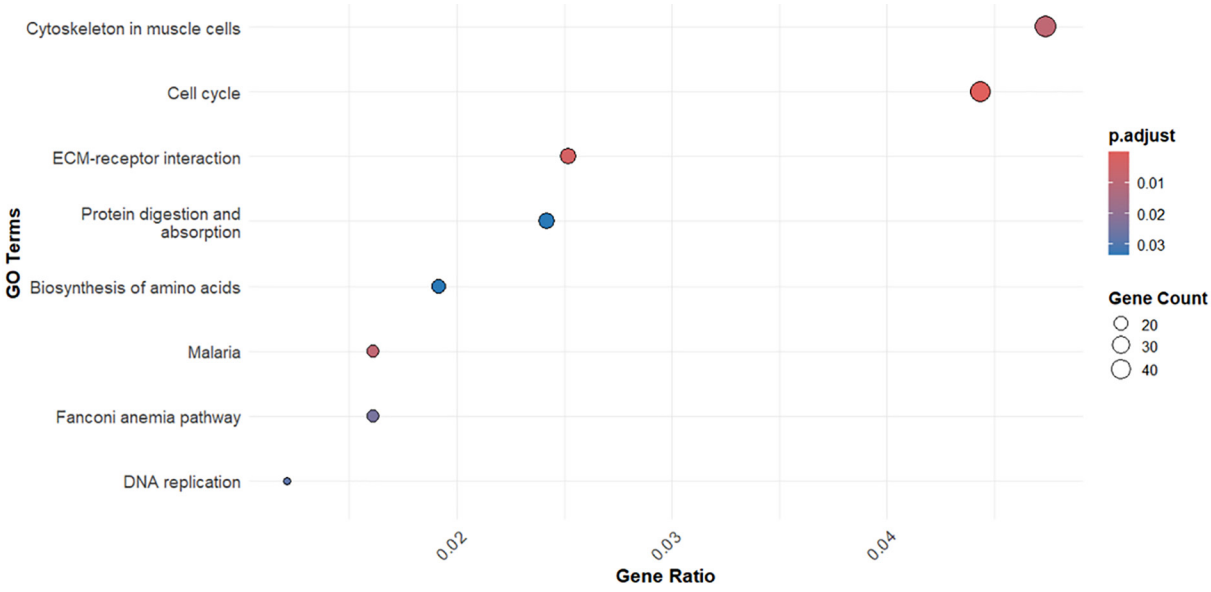


Figure 4. KEGG Pathway Analysis for Normal vs. Tumor samples (N vs. T)

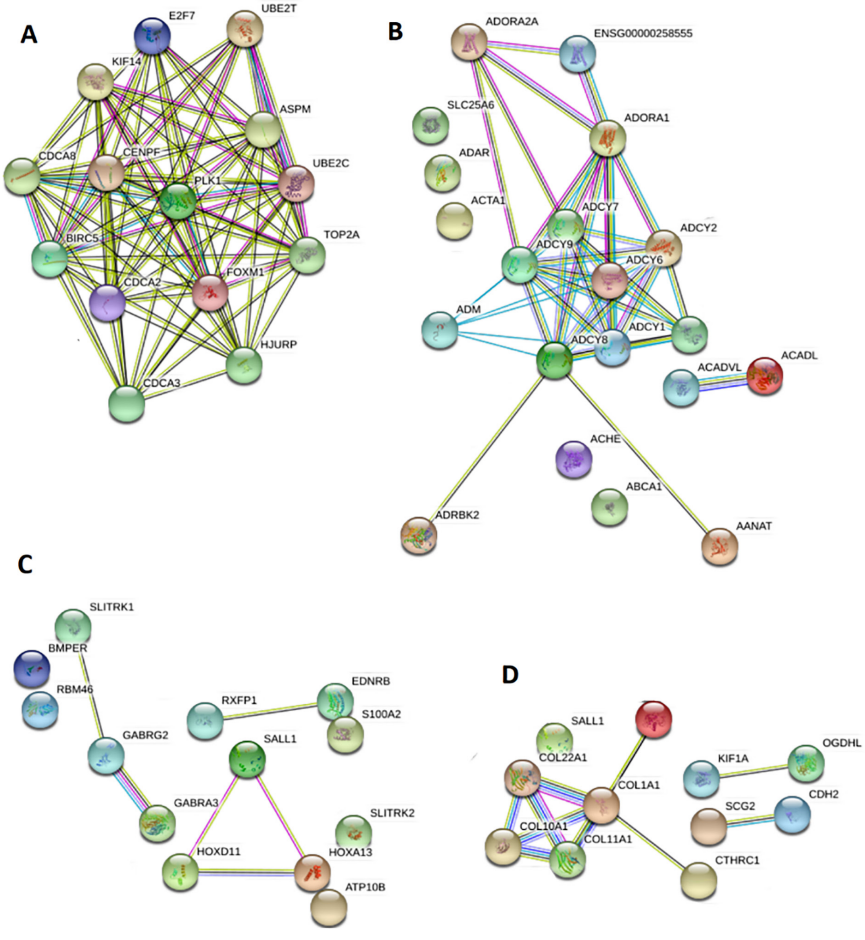


Figure 5. Protein-Protein Interaction (PPI) Networks of Differentially Expressed Genes (DEGs) in Different Sample Comparisons. **A:** Normal vs. Tumor (N vs. T) **B:** Normal Current Smoker vs. Tumor Current Smoker (NC vs. TC) **C:** Normal Former Smoker vs. Tumor Former Smoker (NF vs. TF) **D:** Normal Never Smoker vs. Tumor Never Smoker (NN vs. TN)

Selection of the top 50 genes was based on adjusted p-values (padj), with a confidence score threshold of 400 applied to ensure high reliability in the predicted interactions.

Discussion

The principal objective of this study was to elucidate the molecular pathways through which cigarette smoke influences angiogenesis in NSCLC, utilizing gene expression profiling. While prior research has revealed mechanisms including the upregulation of VEGF, activation of HIF-1α, and induction of inflammatory cytokines (28, 29), our focus on miR-1 was driven by both a biologically grounded hypothesis and the foundational design of the selected dataset. Differential gene expression analysis between normal (N) and tumor (T) tissues identified 2,449 DEGs, several of which play pivotal roles in angiogenesis.

Notably, VEGFC—a critical regulator of angiogenesis and lymphangiogenesis—is significantly upregulated in NSCLC (30, 31) and is likely targeted by miR-1, thereby promoting neovascularization. Similarly, Fibroblast Growth Factor 2 (FGF2), which enhances endothelial cell proliferation and migration (32), and EPHB2, a gene implicated in endothelial cell migration and vascular morphogenesis (33), also represent probable targets. Transforming Growth Factor Beta Receptor 3 (TGFBR3), which modulates TGF-β signaling, is likewise involved in angiogenesis (34).

SPP1 (Osteopontin), a secreted phosphoprotein associated with both tumor progression and angiogenesis (35), appears to be under miR-1 regulation as well. Additionally, Angiopoietin 1 (ANGPT1), a gene crucial for vascular stability and angiogenic signaling (36), and miR-1–regulated genes such as EPHA10 and Fatty Acid Binding Protein 3 (FABP3), involved in Eph receptor signaling and lipid metabolism respectively (37, 38), further underscore the breadth of miR-1's regulatory influence.

In summary, our findings highlight miR-1 as a key regulator of angiogenesis-related genes in NSCLC in response to cigarette smoke exposure.

Comparison of the NF vs. TF samples revealed 697 DEGs, among which genes such as MET, EPAS1, FLT1, TEK, and ANGPT1 showed marked expression differences.

These genes are well-established contributors to angiogenesis and tumor progression (39–41), and likely targets of miR-1, which has been reported to function as a tumor suppressor in NSCLC (42). In the NC vs. TC comparison, which yielded 172 DEGs, angiogenesis-related genes such as *FGF10* and *APLN* were also significantly altered (43, 44), further supporting miR-1's inhibitory role in vascular development.

The NN vs. TN comparison identified 94 DEGs, including *COL1A1*, *COL3A1*, and *BMPER*, all of which are associated with extracellular matrix remodeling and angiogenesis (45). Once again, these genes represent plausible miR-1 targets, underscoring its anti-angiogenic function via gene repression.

GSEA revealed that the N vs. T comparison exhibited significant enrichment in biological processes related to chromosome segregation and keratinization, suggesting widespread alterations in cell cycle regulation and structural integrity in NSCLC. Importantly, GSEA outcomes varied across comparisons (NF vs. TF, NC vs. TC, and NN vs. TN), reflecting the intrinsic heterogeneity of tumor biology and emphasizing that results derived from one tissue context cannot be extrapolated to another without careful consideration.

Consistent with GSEA, GO enrichment analysis of the N vs. T dataset revealed significant enrichment in mitosis- and cell cycle–related terms, including "mitotic nuclear division" and "chromosome segregation."

These findings, illustrated by high Gene Ratios and low adjusted p-values in Figure 4, highlight the functional relevance of the DEGs identified, suggesting their involvement in

maintaining genomic stability and promoting uncontrolled proliferation in NSCLC.

KEGG pathway enrichment analysis further supported these findings, identifying cell cycle and DNA replication pathways as significantly dysregulated. Of particular interest is the enrichment of the Fanconi anemia pathway, implicating defective DNA repair in NSCLC tumorigenesis—given that dysfunction in the FA/BRCA pathway underpins genomic instability and impaired DNA damage response (46). Differences in KEGG enrichment patterns across sample comparisons reflect the diverse molecular landscapes characteristic of lung tumors.

The PPI network derived from N vs. T DEGs revealed robust interactions among several hub genes. Proteins such as *E2F7*, *UBE2T*, and *PLK1* were found to play central roles in cell cycle regulation and the promotion of cellular proliferation (47–49). Their upregulation in NSCLC suggests potential utility as diagnostic biomarkers or therapeutic targets. A stringent interaction confidence score of ≥400 was applied to ensure the reliability of these associations. Notably, interactions among genes such as *FOXM1*, *TOP2A*, and *BIRC5* suggest the existence of a tightly coordinated network that facilitates tumor growth and progression. Cancer cells exhibit a unique capacity to induce angiogenesis, thereby securing an increased supply of oxygen and nutrients, primarily through the secretion of vascular endothelial growth factor (VEGF). The DEGs identified in this study likely contribute to this pro-angiogenic phenotype by driving both cellular proliferation and genomic instability.

For instance, rapid tumor growth and elevated cell density often induce hypoxic conditions, which in turn activate VEGF and other angiogenic mediators (50).

Nevertheless, this study is not without limitations, including a relatively small sample size and reliance on a single dataset. Several other studies have explored the molecular consequences of cigarette smoke exposure in lung cancer. For example, de Biase et al. (2024) analyzed transcriptomic data from nasal and bronchial epithelial samples to identify DEGs and pathway-level alterations linked to lung cancer risk. While their study utilized computational modeling and transcriptional regulatory networks, our work emphasized GSEA, GO, and PPI network analyses (51). Takyar et al. also investigated the effects of cigarette smoke on endothelial cell function, highlighting miR-1 degradation via the VEGF–PI3K–AKT signaling axis and exploring both cellular and clinical mechanisms in depth (52). Additional studies have corroborated our findings, implicating smoking in significant transcriptional reprogramming associated with lung carcinogenesis. One such investigation underscored the role of cell cycle genes, including *E2F* and *PLK1*, and DNA repair pathways, using data from TCGA and GEO, and further examined sex-based differences in epithelial-mesenchymal transition (EMT)-related gene expression (53).

The present study aligns with the findings of Beane et al., who employed RNA sequencing to examine gene expression alterations induced by smoking in the airway epithelium. Their research demonstrated that smoking activates pathways involved in xenobiotic metabolism, oxidoreductase activity, and inflammatory responses, including chemotaxis signaling and cytokine receptor pathways, in both smokers and lung cancer patients. By contrast, our study specifically focused on the role of miR-1 and its downstream targets—VEGFC, FGF2, and ANGPT1. We demonstrated that smoking not only stimulates canonical inflammatory pathways but also exerts a profound effect on angiogenic signaling by downregulating miR-1 expression (54). In summary, prior research, including the review conducted by Prtty et al., underscores the potential of gene expression profiling in NSCLC as a platform for identifying diagnostic biomarkers and

therapeutic targets. Their work emphasized the utility of transcriptomic data in subclassifying tumors, predicting clinical outcomes, and guiding personalized treatment strategies (55). Building upon these foundational insights, our study contributes functional evidence regarding the impact of cigarette smoke on the NSCLC transcriptome, thereby uncovering novel avenues for therapeutic intervention (Figure 6).

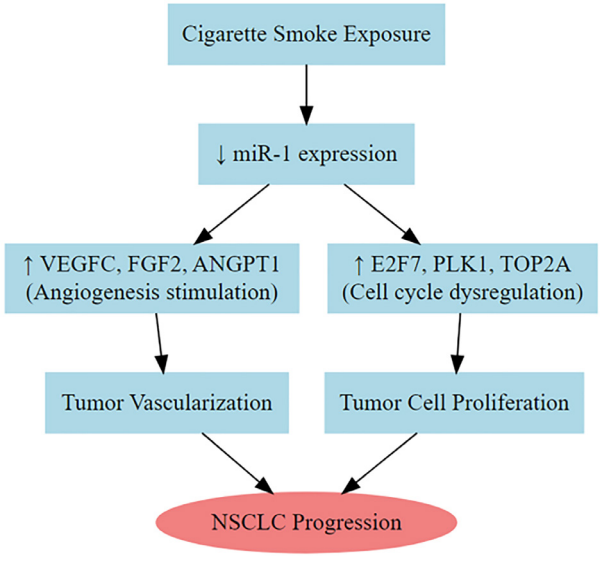


Figure 6. Proposed Mechanism of Cigarette Smoke-Induced NSCLC Progression through Downregulation of miR-1 and Activation of Cell Cycle and Angiogenic Pathways

To enhance the reliability of future investigations, it is imperative to incorporate larger sample sizes and a broader range of datasets. Furthermore, experimental validation through functional assays is essential to substantiate the bioinformatic predictions. It is important to note that although the identified DEGs offer valuable insights, they do not fully elucidate the precise molecular mechanisms underlying tumorigenesis. Thus, continued research in this domain remains critically necessary.

Conclusion

This study aimed to elucidate the impact of cigarette smoke on angiogenesis in NSCLC. We systematically filtered and analyzed key genes regulated by miR-1 among DEGs. Transcriptomic profiling revealed DEGs associated with angiogenesis, cell cycle progression, and DNA repair, thereby highlighting the molecular heterogeneity characteristic of NSCLC. We identified several pivotal genes and hub regulators—such as E2F7 and PLK1—that may serve as potential biomarkers for the diagnosis and targeted treatment of NSCLC. Based on these findings, we propose that therapeutic inhibition of miR-1 and its downstream effectors could represent a promising strategy for managing various cancers, particularly NSCLC. Nonetheless, in order to strengthen the robustness and clinical relevance of these observations, subsequent studies should adopt a more comprehensive approach—utilizing larger datasets and incorporating experimental validation—to further explore the therapeutic and diagnostic implications of miR-1-mediated regulation.

Acknowledgments

We gratefully acknowledge the GEO database for providing access to the dataset (GSE290190) utilized in this study.

Conflict of Interest

The authors declare no conflict of interest.

Funding

No external funding was received for this research.

Ethical Consideration

Ethical approval was not required for this study, as all analyses were conducted in accordance with ethical standards governing the use of publicly available datasets.

Authors Contribution

Kianoush Mohammadi conducted the data analyses and drafted the manuscript. Reza

Safaralizadeh conceptualized and supervised the study and provided critical revisions. Both authors reviewed and approved the final version of the manuscript.

References

1 Reddy R, Reddy S. Trends in imaging patterns of bronchogenic carcinoma: reality or a statistical variation? A single-center cross-sectional analysis of outcomes. Med Princ Pract. 2022; 31 (5): 480-485.
2 Aokage K, Yoshida J, Hishida T, Tsuboi M, Saji H, Okada M, et al. Limited resection for early-stage non-small cell lung cancer as function-preserving radical surgery: a review. Jpn J Clin Oncol. 2017; 47 (1): 7-11.
3 Chen J. A comparative analysis of lung cancer incidence and tobacco consumption in Canada, Norway and Sweden: a population-based study. Int J Environ Res Public Health. 2023; 20 (20): 6930.
4 Xie S, Wu Z, Qi Y, Wu B, Zhu X. The metastasizing mechanisms of lung cancer: recent advances and therapeutic challenges. Biomed Pharmacother. 2021; 138: 111450.
5 Piñeros M, Parkin DM, Ward K, Chokunonga E, Ervik M, Farrugia H, et al. Essential TNM: a registry tool to reduce gaps in cancer staging information. Lancet Oncol. 2019; 20 (2): e103-e111.
6 Yang D, Dang S, Wang Z, Xie M, Li X, Ding X. Vessel co-option: a unique vascular-immune niche in liver cancer. Front Oncol. 2024; 14: 1386772.
7 Sobus SL, Warren GW. The biologic effects of cigarette smoke on cancer cells. Cancer. 2014; 120 (23): 3617-3626.
8 Korde A, Jin L, Zhang JG, Ramaswamy A, Hu B, Kolahian S, et al. Lung endothelial microRNA-1 regulates tumor growth and angiogenesis. Am J Respir Crit Care Med. 2017; 196 (11): 1443-1455.
9 Korde A, Ramaswamy A, Anderson S, Hu B, Velasco Torrez W, Pisani MA, et al. Cigarette smoke activates the angiogenic switch in the lung cancer field by downregulating microRNA-1 [abstract]. Am J Respir Crit Care Med. 2023; 207: A6530.
10 Hua L, Zhu G, Wei J. MicroRNA-1 overexpression increases chemosensitivity of non-small cell lung cancer cells by inhibiting autophagy related 3-mediated autophagy. Cell Biol Int. 2018; 42 (9): 1240-1249.
11 Condrat CE, Thompson DC, Barbu MG, Bugnar OL, Boboc A, Cretoiu D, et al. miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis. Cells. 2020; 9 (2): 276.
12 Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. R package version 1.46.0. 2023. Available from: https://bioconductor.org/packages/DESeq2
13 Wickham H. ggplot2: Elegant graphics for data analysis. R package version 3.5.1. 2024. Available from: https://ggplot2.tidyverse.org
14 Blighe K, Rana S, Lewis M. EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling. R package version 1.24.0. 2023. Available from: https://github.com/kevinblighe/EnhancedVolcano
15 Kolde R. pheatmap: Pretty heatmaps. R package version 1.0.12. 2019. Available from: https://CRAN.R-project.org/package=pheatmap
16 Wickham H, François R, Henry L, Müller K, Vaughan D. dplyr: A grammar of data manipulation. R package version 1.1.4. 2023. Available from: https://CRAN.R-project.org/package=dplyr
17 Wickham H, Averick M, Bryan J, Chang W, McGowan L, François R, et al. Welcome to the tidyverse. J Open Source Softw. 2023; 8 (86): 1686.
18 Carlson M. org.Hs.eg.db: Genome wide annotation for Human. R package version 3.20.0. 2023.
19 Szklarczyk D, Gable AL, Lyon D, Junge A, Wyder S, Huerta-Cepas J, et al. STRING v11: protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. Nucleic Acids Res. 2019; 47 (D1): D607-D613.
20 Maza E. Comparison of TMM (edgeR), RLE (DESeq2), and MRN normalization methods for a simple two-conditions-without-replicates RNA-Seq experimental design. Front Genet. 2016; 7: 164.
21 Llambrich M, Correig E, Gumà J, Brezmes J, Cumeras R. Amanida: an R package for meta-analysis of metabolomics non-integral data. Bioinformatics. 2022; 38 (2): 583-585.
22 van Lingen HJ, Suarez-Diez M, Saccenti E. Normalization of gene counts affects principal components-based exploratory analysis of RNA-sequencing data. BBA Gene Regul Mech. 2024; 1867 (4): 195058.
23 Liu S, Xie X, Lei H, Zou B, Xie L. Identification of key circRNAs/lncRNAs/miRNAs/mRNAs and pathways in preeclampsia using bioinformatics analysis. Med Sci Monit. 2019; 25: 1679-1693.
24 Chawla K, Kuiper M. Genes2GO: a web application for querying gene sets for specific GO terms. Bioinformation. 2016; 12 (3): 231-232.
25 Yan ZC, Liu TH, Yu SB, Zhan YH, Wang ZY, Zhu ZW, et al. Expression and significance analysis of GDF3 in testicular cancer based on TCGA and

GTEx databases. Zhonghua Nan Ke Xue. 2023; 29 (12): 980-985.
26 Trotta RJ, Swanson KC, Klotz JL, Harmon DL. Influence of postruminal casein infusion and exogenous glucagon-like peptide 2 administration on the jejunal mucosal transcriptome in cattle. PLoS One. 2024; 19 (8): e0308983.
27 Mortazavi A, Ghaderi-Zefrehei M, Muhaghegh Dolatabady M, Golshan M, Nazari S, Sadr AS, et al. An integrated bioinformatics approach to identify network-derived hub genes in starving zebrafish. Animals (Basel). 2022; 12 (19): 2724.
28 Ejaz S, Lim CW. Toxicological overview of cigarette smoking on angiogenesis. Environ Toxicol Pharmacol. 2005; 20 (2): 335-344.
29 Ao L, Gao H, Jia L, Liu S, Guo J, Liu B, et al. Matrine inhibits synovial angiogenesis in collagen-induced arthritis rats by regulating HIF-VEGF-Ang and inhibiting the PI3K/Akt signaling pathway. Mol Immunol. 2022; 141: 13-20.
30 Lymboussaki A, Olofsson B, Eriksson U, Alitalo K. Vascular endothelial growth factor (VEGF) and VEGF-C show overlapping binding sites in embryonic endothelia and distinct sites in differentiated adult endothelia. Circ Res. 1999; 85 (11): 992-999.
31 Li J, Hong M, Pan T. Clinical significance of VEGF-C and VEGFR-3 expression in non-small cell lung cancer. J Huazhong Univ Sci Technolog Med Sci. 2006; 26 (5): 587-590.
32 Seghezzi G, Patel S, Ren CJ, Gualandris A, Pintucci G, Robbins ES, et al. Fibroblast growth factor-2 (FGF-2) induces vascular endothelial growth factor (VEGF) expression in the endothelial cells of forming capillaries: an autocrine mechanism contributing to angiogenesis. J Cell Biol. 1998; 141 (7): 1659-1673.
33 Salvucci O, Tosato G. Essential roles of EphB receptors and EphrinB ligands in endothelial cell function and angiogenesis. Adv Cancer Res. 2012; 114: 21-57.
34 Deng X, Ma N, He J, Xu F, Zou G. The role of TGFBR3 in the development of lung cancer. Protein Pept Lett. 2024; 31 (7): 491-503.
35 Pang X, Gong K, Zhang X, Wu S, Cui Y, Qian BZ. Osteopontin as a multifaceted driver of bone metastasis and drug resistance. Pharmacol Res. 2019; 144: 235-244.
36 Bilimoria J, Singh H. The angiopoietin ligands and Tie receptors: potential diagnostic biomarkers of vascular disease. J Recept Signal Transduct Res. 2019; 39 (3): 187-193.
37 Huang S, Ma L, Liu X, He C, Li J, Hu Z, et al. A non-coding variant in 5' untranslated region drove up-regulation of pseudo-kinase EPHA10 and caused non-syndromic hearing loss in humans. Hum Mol Genet. 2023; 32 (5): 720-731.
38 Li B, Syed MH, Khan H, Singh KK, Qadura M. The role of fatty acid binding protein 3 in cardiovascular diseases. Biomedicines. 2022; 10 (9): 2283.
39 Vlachostergios PJ, Tamposis IA, Anagnostou M, Papathanassiou M, Mitrakas L, Zachos I, et al. Hypoxia-inducible factor-2-altered urothelial carcinoma: clinical and genomic features. Curr Oncol. 2022; 29 (11): 8638-8649.
40 Kurmyshkina O, Kovchur P, Schegoleva L, Volkova T. Markers of angiogenesis, lymphangiogenesis, and epithelial-mesenchymal transition (plasticity) in CIN and early invasive carcinoma of the cervix: exploring putative molecular mechanisms involved in early tumor invasion. Int J Mol Sci. 2020; 21 (18): 6515.
41 Zhao Y, Yu B, Wang Y, Tan S, Xu Q, Wang Z, et al. Ang-1 and VEGF: central regulators of angiogenesis. Mol Cell Biochem. 2025; 480 (2): 621-637.
42 Han C, Yu Z, Duan Z, Kan Q. Role of microRNA-1 in human cancer and its therapeutic potentials. Biomed Res Int. 2014; 2014 : 428371.
43 Liu F, Li G, Deng L, Kuang B, Li X. The roles of FGF10 in vasculogenesis and angiogenesis. Biomed Res. 2017; 28 (3): 1329-1332.
44 Lugano R, Ramachandran M, Dimberg A. Tumor angiogenesis: causes, consequences, challenges and opportunities. Cell Mol Life Sci. 2020; 77 (9): 1745-1770.
45 Pankratz F, Maksudova A, Goesele R, Meier L, Proelss K, Marenne K, et al. BMPER Improves Vascular Remodeling and the Contractile Vascular SMC Phenotype. Int J Mol Sci. 2023; 24 (5): 4950.
46 Yao C, Du W, Chen H, Xiao S, Huang L, Chen F. The Fanconi anemia/BRCA pathway is involved in DNA interstrand cross-link repair of adriamycin-resistant leukemia cells. Leuk Lymphoma. 2015; 56 (3): 755-762.
47 Di Stefano L, Jensen MR, Helin K. E2F7, a novel E2F featuring DP-independent repression of a subset of E2F-regulated genes. EMBO J. 2003; 22 (23): 6289-6298.
48 Liu LL, Zhu JM, Yu XN, Zhu HR, Shi X, Bilegsaikhan E, et al. UBE2T promotes proliferation via G2/M checkpoint in hepatocellular carcinoma. Cancer Manag Res. 2019; 11: 8359-8370.
49 Iliaki S, Beyaert R, Afonina IS. Polo-like kinase 1 (PLK1) signaling in cancer and beyond. Biochem Pharmacol. 2021; 193: 114747.
50 Saman H, Raza SS, Uddin S, Rasul K. Inducing

angiogenesis, a key step in cancer vascularization, and treatment approaches. Cancers (Basel). 2020; 12 (5): 1172.

51 de Biase MS, Massip F, Wei TT, Giorgi FM, Stark R, Stone A, et al. Smoking-associated gene expression alterations in nasal epithelium reveal immune impairment linked to lung cancer risk. Genome Med. 2024; 16 (1): 54.

52 Korde A, Ramaswamy A, Anderson S, Jin L, Zhang JG, Hu B, et al. Cigarette smoke induces angiogenic activation in the cancer field through dysregulation of an endothelial microRNA. Commun Biol. 2025; 8: 511.

53 Hammouz RY, Kostanek JK, Dudzisz A, Witas P, Orzechowska M, Bednarek AK. Differential expression of lung adenocarcinoma transcriptome with signature of tobacco exposure. J Appl Genet. 2020; 61 (3): 421-437.

54 Beane J, Vick J, Schembri F, Anderlind C, Gower A, Campbell J, et al. Characterizing the impact of smoking and lung cancer on the airway transcriptome using RNA-Seq. Cancer Prev Res (Phila). 2011; 4 (6): 803-817.

55 Petty RD, Nicolson MC, Kerr KM, Collie-Duguid E, Murray GI. Gene expression profiling in non-small cell lung cancer: from molecular mechanisms to clinical application. Clin Cancer Res. 2004; 10 (10): 3237-3248.