Contents lists available at *Dergipark*

Journal of Scientific Reports-A

journal homepage: *https://dergipark.org.tr/tr/pub/jsr-a*

E-ISSN: 2687-6167 ***Number 60, March 2025***

RESEARCH ARTICLE

Receive Date: 08.11.2024 *Accepted Date: 23.01.2025*

MiR-3605, miR-511 and miR-6788: Potential diagnostic and prognostic biomarkers for squamous cell carcinoma of the lung

Medi Kori[a*]

[a]*Department of Medical Biotechnology, Institute of Health Sciences, Acibadem Mehmet Ali Aydinlar University, Istanbul 34718, Türkiye, ORCID: 0000-0002-4589-930X*

Abstract

Squamous cell carcinoma of the lung (LUSC) is the second most common subtype of lung cancer and lung cancer is responsible for most cancer-related deaths. It can therefore be assumed that there is still a large gap in reducing a significant proportion of the global cancer burden. To close this gap, new methods are needed that provide better diagnostic and prognostic approaches for LUSC. Given the advantages of miRNA biomolecules as potential biomarkers, a systems biology approach was used in this study to define diagnostic and/or prognostic miRNA biomarker candidates for LUSC. Accordingly, the differentially expressed genes (DEGs) of LUSC were identified by processing RNA-Seq expression data. After analyzing the DEGs, a reporter feature algorithm was applied, which yielded reporter miRNAs that have significant potential as biomarker candidates. Using miRNA-Seq data from LUSC, the potential diagnostic and prognostic performance of reporter miRNAs was investigated. Using this approach, miR-3605 and miR-6788 were found to have diagnostic capabilities in LUSC, while miR-511, which was found in serum, had diagnostic and prognostic properties. Overall, this study offers precious data for further experimental and clinical efforts to diagnose and predict LUSC, and the presented diagnostic and/or prognostic miRNAs were associated with LUSC for the first time in this study.

Keywords: Squamous cell carcinoma of the lung (LUSC); miRNA biomarkers; diagnosis and prognosis; miR-511; miR-3605; miR-6788.

* Corresponding author. Tel.: +90-216-500-4196
E-mail address: medi.kori@acibadem.edu.tr

1. Introduction

Cancer is an unrestrained cellular division process that leads to cell alteration and is responsible for most causes of death worldwide. Among the various types of cancer, lung cancer accounts for a large proportion of cancer deaths according to cancer statistics [1]. Histopathologically, lung cancer is divided into 2 main types, small cell lung cancer (SCLC) and non-small cell lung cancer (NSCL), and among these two types, the incidence of NSLC is much higher compared to SCLC. In addition, non-small cell lung cancers are categorized into different subtypes, including lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) [2].

The statistics on the incidence and high mortality rate show that we need new strategies for the diagnosis, prevention and treatment of NSLCs, and as LUSC cases fade into the background, any innovation could open up new avenues in the clinic. Biomarkers have great potential to fulfil the requirements for a better diagnosis or prognosis of a disease. According to the BEST glossary, a biomarker is "a characteristic that is measured as an indicator of normal biological processes, pathogenic processes or responses to an exposure or intervention, including therapeutic interventions" [3].

Given the great potential of miRNAs as biomarker candidates for disease diagnosis and prognosis [4], a novel systems biology approach was adopted in the present study to identify miRNA biomarkers that can be used for LUSC for diagnostic and prognostic purposes. Accordingly, the differentially expressed genes (DEGs) of LUSC were determined by processing RNA-Seq expression data. The obtained DEGs were integrated into a comprehensive human miRNA gene-target network and reporter miRNAs representing the major transcriptional changes were identified. The diagnostic and prognostic potential of reporter miRNAs was evaluated using miRNA-Seq expression data (Figure 1). Consequently, diagnostic and/or prognostic miRNA candidates that could be considered as diagnostic and prognostic biomarkers in LUSC were presented in this study.

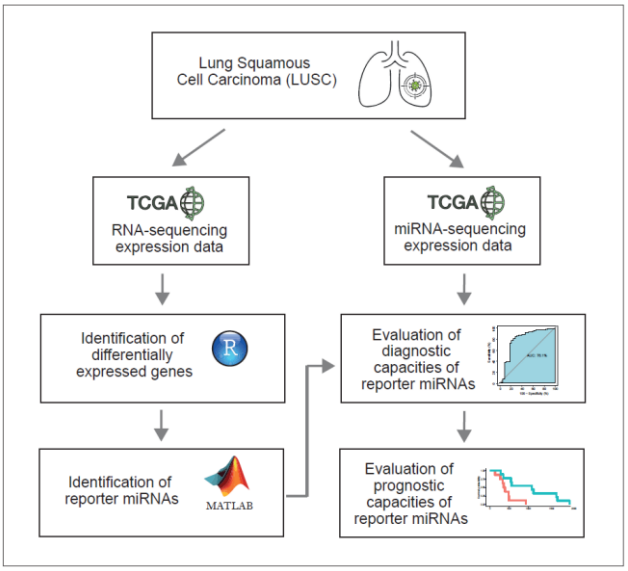

Fig. 1. The overall methodology used in the study.

2. Materials and Methods

2.1. RNA-seq and miRNA-seq datasets of squamous cell carcinoma of the lung

The RNA-Seq data and miRNA-Seq data from LUSC were obtained from The Cancer Genome Atlas (TCGA-LUSC) [5]. The RNA-Seq data included 502 primary solid tumors and 49 normal subjects, while the miRNA-Seq data included 478 primary solid tumors and 45 normal subjects (Figure 2A). The RNA-Seq data were used for the identification of reporter miRNAs, and the miRNA-Seq data were used for diagnostic analyzes of reporter miRNAs.

2.2. Identification of differentially expressed genes

The DESeq2 package [6] in R/Bioconductor (version 4.0.2) [7] was used to determine DEGs between LUSC and normal phenotypes. DESeq2 fits a generalized linear model from the family of negative binomial models and uses the Wald test. The Benjamini-Hochberg method was used to control for false discovery rate (FDR). An adjusted p-value of $< 1\times10^{-15}$ was used as the statistical significance threshold. Fold change (FC) was used to assess the expression pattern of DEGs (i.e., up- or down-regulation), and at least an FC of 2 was considered significant.

2.3. Identification of reporter miRNAs

Reporter miRNAs were determined using the reporter features algorithm [8], which was previously fitted and used in several studies [9-12]. The fitted algorithm was performed in MATLAB (R2016) in the present study.

The original algorithm aimed to identify reporter metabolites by integrating differential transcriptional data (in the form of p-values representing significant changes in gene expression between different phenotypes) with a metabolic model (consisting of gene-response-metabolite interactions). Subsequently, in this study, the miRNA-target gene network was integrated with genes that differ significantly between LUSC and normal phenotypes (i.e. DEGs). The miRNA-target gene network was reconstructed using the experimentally validated miRNA-target gene interactions [13] and miRTarbase (version 6.0) [14].

By applying the adapted algorithm, the p-values of the genes were converted into z-scores using the inverse cumulative distribution and considering the interactions between miRNA and target DEGs. Accordingly, each miRNA was assigned a z-score and then the values that followed a standard normal distribution were converted to p-values. In this study, statistically significant miRNAs with a p-value < 0.001 were classified as reporter miRNAs.

2.4. Diagnostic validation of reporter miRNAs

The miRNA-Seq expression data from TCGA were used to evaluate the diagnostic performance of reporter miRNAs. Sensitivity (the proportion of how well the test identifies true positives) and specificity (the proportion of how well the test identifies true negatives) were used to predict the diagnostic ability of reporter miRNAs based on the Receiver Operating Characteristic Curve. The area under the Receiver Operating Characteristic curve (AUC) was computed to assess the diagnostic ability of the reporter miRNAs, and reporters with an AUC value $\geqslant 0.75$ were considered diagnostic reporter candidates [15].

2.5. Prognostic validation of reporter miRNAs

To evaluate the prognostic capabilities of the reporter miRNAs, the clinical information of the LUSC samples was extracted from the miRNA-Seq data of TCGA and used for prognostic performance analysis. The prognostic performance analyses were performed for each reporter. Survival analyses were performed to evaluate the prognostic capacity of the reporter miRNAs. In the analysis, samples were categorized into two groups (high- and low- risk)

based on the clinical information of the samples (i.e. prognostic index - PI), where PI is the linear component of the Cox model. Survival signatures were assessed using Kaplan-Meier (K-M) plots and log-rank p-values. The hazard ratio (HR) was computed using the ratio between the relative mortality rates of group 1 and group 2 ($HR = (O_1/E_1)/(O_2/E_2)$, where O refers to the observed number of deaths and E refers to the expected number of deaths). Reporter miRNAs with a log-rank p-value < 0.05 were assumed to have prognostic abilities.

2.6. Biological interpretation of reporter miRNAs

To gain biological insights into reporter miRNAs, miRBase: the microRNA database the archive for microRNA sequences and annotations repository [16] was used. As knowledge of extracellular circulating miRNAs is crucial in the clinic [17], the miRandola (Extracellular Circulating MicroRNAs Database) resource [18] was also used to determine whether reporter miRNAs are present in extracellular body fluids such as serum, plasma, saliva and urine.

3. Results

3.1. Differentially expressed genes and reporter miRNAs of squamous cell carcinoma of the lung

The obtained RNA-Seq expression data were used to determine DEGs between LUSC and normal samples. According to the statistical significance considered in this study, a total of 4271 up-regulated and 2787 down-regulated DEGs were determined (Figure 2B). The DEGs and their adjusted p-values were used as proxies for the determination of reporter miRNAs. As a result of implementing the reporter features algorithm, a total of 13 reporter miRNAs were determined with a p value of < 0.001. Among the resulting reporter miRNAs, the most statistically significant reporters were miR-768 (p-value = 1×10^{-15}) and miR-3605 (p-value = 1.89×10^{-7}) (Figure 2C).

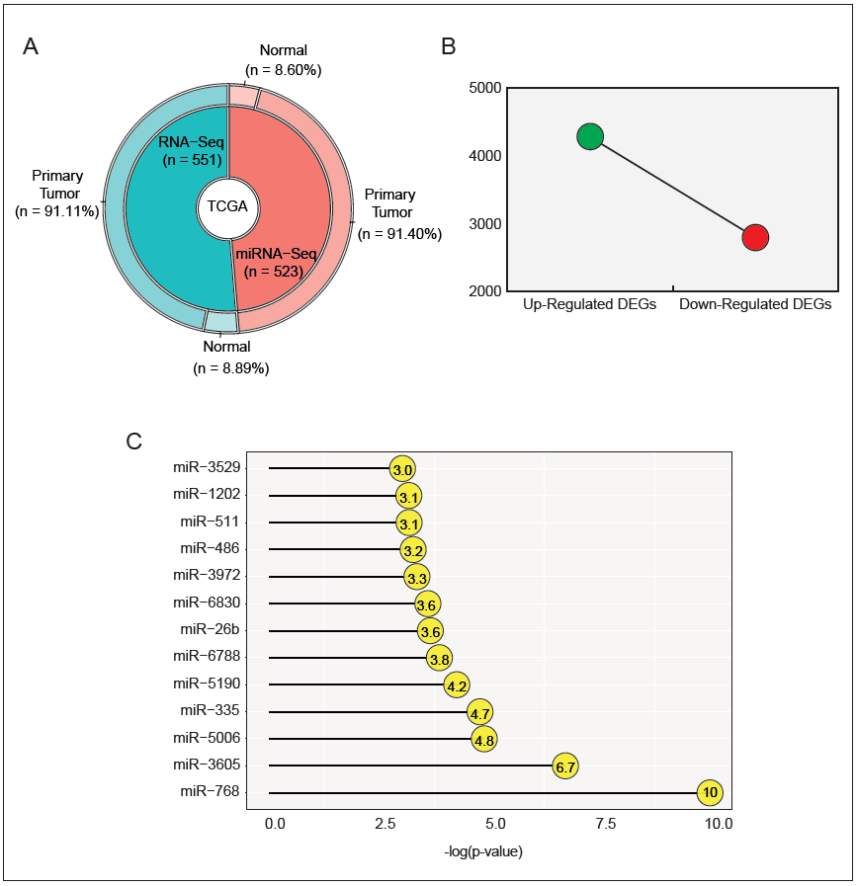

Fig. 2. The number of samples of DEGs and reporter miRNAs. (A) The number of RNA and miRNA sequencing data of primary tumor and control samples from the TCGA-LUSC. (B) The number of DEGs identified during the analysis. (C) The distribution of p-values of reporter miRNAs.

3.2. Diagnostic power of reporter miRNAs

In order to investigate the diagnostic capacities of the culminated reporter miRNAs, miRNA-seq expression data from TCGA-LUSC were used. Of the 13 discovered reporter miRNAs, miRNA-seq expression data are not available for two miRNAs (miR-486 and miR-768). Therefore, these reporter miRNAs cannot be included in the analysis of diagnostic performance analyses. In the diagnostic performance analysis, each reporter miRNA was evaluated based on the AUC values of the Receiver Operating Characteristic curves.

The diagnostic analyzes of the reporter miRNAs showed that the three reporter miRNAs had high diagnostic capacity (AUC value ≥ 0.75). The differences in miRNA expression values between the cancer samples and the healthy subjects were represented by box plots for these three diagnostic miRNAs considering their p-values (Figure 3). The obtained AUC values for the reporters are as follows: for miR-3605, the AUC value is 0.786 (Figure 3A); for miR-511 (Fig 3B), the AUC value is 0.867; and for miR-6788 (Figure 3C), the AUC value is 0.793.

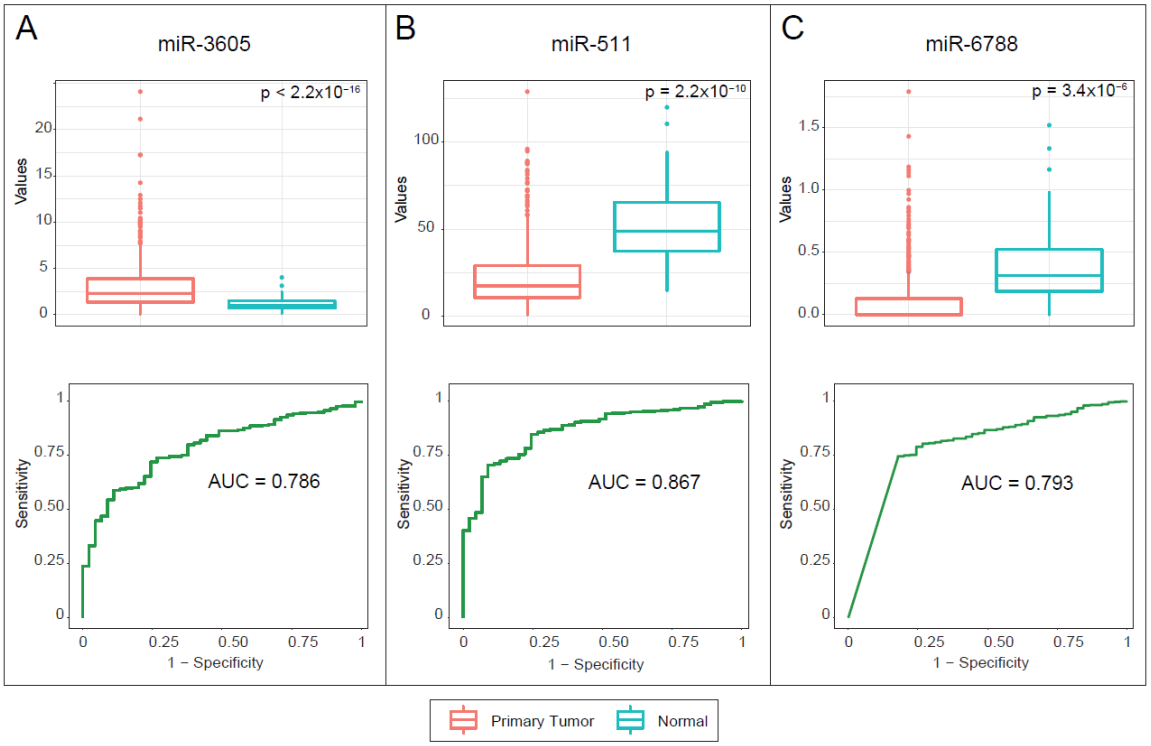

Fig. 3. The diagnostic capabilities of reporter miRNAs. Boxplots show the expression levels of reporter miRNAs between primary tumor and control groups with p-values. The Receiver Operating Characteristic curves show the AUC values of the individual reporters. (A) miR-3605. (B) miR-511. (C) miR-6788.

3.3. Prognostic power of reporter miRNAs

Survival data from TCGA-LUSC were used to assess the prognostic ability of reporters. The prognostic ability of the reporter miRNAs was evaluated using K-M plots, taking into account the log-rank p-value and HRs. Accordingly, miR-511 was found to be statistically significant with a log-rank p-value of 0.04 and a HR of 1.32 according to the criteria considered (Figure 4). The other two miRNAs that indicated remarkable diagnostic ability (i.e. miR-3605 and miR-6788) did not show statistical significance in accordance with their log rank p-values (p-value > 0.05). Considering these results, miR-511 was shown to have both high diagnostic and prognostic capacity in LUSC cases.

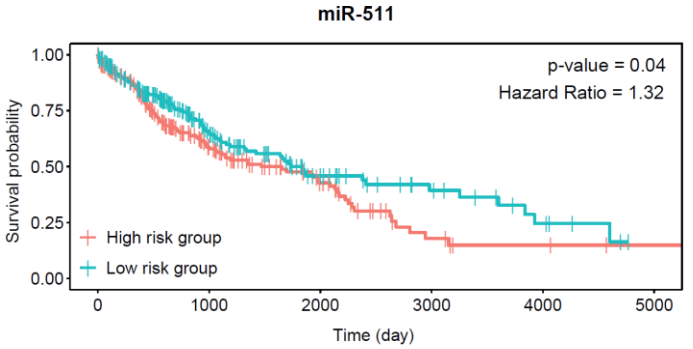

Fig. 4. The prognostic ability of miR-511. K-M diagram to estimate the survival time of LUSC patients as well as log-rank p-value and hazard ratio were shown. The high- and low-risk groups are shown in red and blue respectively.

3.4. Biological features of reporter miRNAs

To investigate the biological properties of the resulting diagnostic and prognostic reporter miRNAs, the miRBase database [16] and the miRandola resource [18] were searched. It was found that 5 of the miRNAs have different biological functions. miR-26a is known for its interaction with genes such as EZH2 and TET1. miR-29a is associated with the delay of atherosclerosis as well as matrix remodeling and fibrosis. miR-205 plays a role in epithelial–mesenchymal transition processes. miR-375 is associated with the differentiation of embryonic stem cells into liver and insulin secretory cells. miR-99b is known for its suppressive effect on signal transduction pathways (i.e. NF-κB pathways) [16].

In addition, miRNAs in different cell types (i.e. cancer or normal cells) can be secreted outside the cells and then enter circulating body fluids (i.e. serum, plasma, saliva, etc.). MiRNAs circulating in body fluids offer great potential in the clinic and are even more advantageous if they are to be used as biomarkers. Accordingly, when screening the potential of culminated reporter miRNAs in body fluids using the miRandola resource [18], it was revealed that the diagnostic and prognostic reporter miR-511 could be present in serum.

4. Discussion

The miRNAs are a group of small single-stranded non-coding RNAs that are involved in regulating the expression of many genes and control remarkable biological processes such as cell division, cell differentiation, cell death, cell migration, angiogenesis and tumorigenesis [19]. The miRNAs can regulate the expression of about 60% of human genes, and today the link between miRNA and cancer, namely the decreased or increased expression of miRNAs in cancer cells, has been proven and recognized by the authorities [20]. Since miRNAs have advantages over other biomolecules in many respects, miRNAs are considered potential markers for cancer. Accordingly, the fact that miRNAs have high biological stability, are present in body fluids and are regulated at different stages of carcinogenesis, brings the use of miRNAs as diagnostic and prognostic biomarkers in the clinic to the forefront [21]. In addition to the potential of miRNAs as biomarkers, miRNAs have also been highlighted for their potential, especially for their success in diagnosis and prognosis in early stages and various stages of diseases. Indeed, miRNAs have a remarkable impact on the diagnosis of even inconspicuous cancers at an early stage and show different survival signatures at different stages of the diseases [22, 23].

In this study, a reporter feature algorithm was adapted to identify reporter miRNAs. DEGs that showed significant changes in expression between LUSC and control conditions were found by analyzing TCGA data and used as input

for the implementation of the algorithm. The resulting 13 reporter miRNAs were analyzed for their diagnostic and prognostic capabilities using miRNA-Seq data. It was found that 2 of the miRNAs (miR-3605 and miR-6788) had statistically significant diagnostic ability. Although there is limited information about miR-3605 and miR-6788 in the literature, these 2 miRNAs have been associated with 2 different diseases in studies. MiR-3605 has been associated with celiac disease, an immune-mediated disease. This association was established on the basis of differential expression between diseased and healthy samples in miRNA sequencing data [24]. In addition, miR-6788 has been associated with gastric cancer. It has been reported that CNALPTC1 is upregulated in patients with gastric cancer and can promote cell proliferation. MiR-6788, one of the targets of CNALPTC1, mediates the progression of gastric cancer through miR-6788 and there is a negative correlation between them [25].

A miRNA circulating in serum, miR-511, has both diagnostic and prognostic capabilities in LUSC cases. The miR-511 has been associated with various types of cancer. For example, miR-511 was found to act as a tumor suppressor in colorectal cancer and was presented as a prognostic biomarker for this disease [26]. It has also been reported that miR-511 can inhibit the proliferation and metastasis of breast cancer cells by downregulating FGF4 expression [27]. Like miR-6788, miR-511 has also been associated with gastric cancer in previous studies. It has been shown that miR-511 can inhibit cell migration, invasion and epithelial–mesenchymal transition in gastric cancer by targeting PAK2 [28, 29]. In 2014, Zhang and coworkers reported that miR-511 promotes apoptosis of radioresistant lung adenocarcinoma cells by inducing BAX, thus introducing miR-511 as a potential therapeutic molecule especially for the treatment of patients with radioresistant lung adenocarcinoma [30].

The lack of experimental validation of the potential miRNA biomarkers with associated LUSC tissue samples or cell lines was the major drawback of the study. Therefore, future clinical studies need to be conducted to evaluate and examine the diagnostic and prognostic capabilities of the biomarkers found. With this in silico study, which can be seen as one of the most important steps in the development of biomarkers, an important step has been taken. However, in order to appeal to a broad medical group, experimental validation is essential.

Overall, the introduced miR-3605 and miR-6788, which have remarkable diagnostic capacity, and circulating miR-511 in serum, which has both diagnostic and prognostic capacity, were associated with the LUSC cases for the first time in this study. The need for novel biomarkers associated with life-threatening human cancers, including LUSC, remains very high. Given this need, this study provides powerful biomarker candidates that merit further experimental and clinical efforts.

Author Contribution

M.K. organized and performed all the analyses and wrote the manuscript.

Acknowledgements

There is no conflict of interest.

References

[1] R. L. Siegel, K. D. Miller, N.S. Wagle, and A. Jemal, “Cancer statistics, 2023,” *CA Cancer J. Clin.*, vol. 73, no. 1, pp. 17-48, Jan. 2023, doi:10.3322/caac.21763.

[2] A. Lahiri, *et al.*, “Lung cancer immunotherapy: progress, pitfalls, and promises,” *Mol. Cancer*, vol. 22, no. 1, Feb. 2023, doi:10.1186/s12943-023-01740-y.

[3] BEST (Biomarkers, EndpointS, and other Tools) Resource. Maryland, 2016.

[4] C.E. Condrat, *et al.*, “miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis”, *Cells*, vol. 9, no. 2, Jan. 2020, doi:10.3390/cells9020276.

[5] K. Tomczak, P. Czerwińska, and M. Wiznerowicz, “The Cancer Genome Atlas (TCGA): an immeasurable source of knowledge”, *Contemp. Oncol. (Pozn)*, vol. 19, no. 1A, pp. A68-A77, 2023, doi:10.5114/wo.2014.47136.

[6] M. I. Love, W. Huber, and S. Anders, "Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2", *Genome Biol.*, vol. 15, no. 12, 2014, doi:10.1186/s13059-014-0550-8.
[7] W. Huber, *et al*., "Orchestrating high-throughput genomic analysis with Bioconductor," *Nat. Methods*, vol. 12, no. 2, pp. 115-121, Feb. 2015, doi:10.1038/nmeth.3252.
[8] K. R. Patil, and J. Nielsen, "Uncovering transcriptional regulation of metabolism by using metabolic network topology", *Proc. Natl. Acad. Sci. U S A*, vol. 102, pp. 2685-9, Feb. 2005, doi:10.1073/pnas.0406811102.
[9] M. Kori, and K.Y. Arga, "Potential biomarkers and therapeutic targets in cervical cancer: Insights from the meta-analysis of transcriptomics data within network biomedicine perspective", *PLoS One*, vol.13, Jul. 2018, doi: 10.1371/journal.pone.0200717.
[10] M. Kori, and E, Gov, "Bioinformatics Prediction and Machine Learning on Gene Expression Data Identifies Novel Gene Candidates in Gastric Cancer", *Genes*, vol. 13 no. 12, Jan. 2022, doi: 10.3390/genes13122233.
[11] N. Kelesoglu, M. Kori, B. Turanli, K. Y. Arga, B. K. Yilmaz, and O. A. Duru, "Acute Myeloid Leukemia: New Multiomics Molecular Signatures and Implications for Systems Medicine Diagnostics and Therapeutics Innovation", *OMICS*, vol. 26, no. 7, p.p. 392-403, Jul. 2022, doi: 10.1089/omi.2022.0051.
[12] M. Kori, D. Cig, K. Y. Arga, and C. Kasavi, "Multiomics Data Integration Identifies New Molecular Signatures for Abdominal Aortic Aneurysm and Aortic Occlusive Disease: Implications for Early Diagnosis, Prognosis, and Therapeutic Targets," *OMICS*, vol. 26, no. 5, p.p. 290-304, May 2022, doi: 10.1089/omi.2022.0021.
[13] E. Gov, and K.Y. Arga, "Interactive cooperation and hierarchical operation of microRNA and transcription factor crosstalk in human transcriptional regulatory network", *IET Syst. Biol.*, vol. 10, p.p. 219–228, Dec. 2016, doi:10.1049/iet-syb.2016.0001.
[14] C. H. Chou, *et al.*, "miRTarBase 2016: updates to the experimentally validated miRNA-target interactions database", *Nucleic Acids Res.*, vol. 44, no. D1, p.p. D239-D247, Jan. 2016, doi:10.1093/nar/gkv1258.
[15] J. N. Mandrekar. "Receiver operating characteristic curve in diagnostic test assessment", *J. Thorac. Oncol.*, vol. 5, no. 9, p.p. 1315-1316, Sep. 2010, doi: 10.1097/JTO.0b013e3181ec173d.
[16] A. Kozomara, M. Birgaoanu, and S. Griffiths-Jones, "miRBase: from microRNA sequences to function", *Nucleic Acids Res.*, vol. 47, no. D1, p.p. D155–D162, Jan 2019, doi: 10.1093/nar/gky1141.
[17] M. Kori, E. Gov, K. Y. Arga, R. Sinha, "Biomarkers From Discovery to Clinical Application: In Silico Pre-Clinical Validation Approach in the Face of Lung Cancer", *Biomark. Insights*, vol. 3, no. 19, Oct. 2024, doi:10.1177/11772719241287400.
[18] F. Russo, *et al.*, "miRandola: extracellular circulating microRNAs database", *PLoS One*, vol. 7, no. 10, 2012, doi:10.1371/journal.pone.0047786.
[19] R. Karaismailoglu, and S. Marakli, "miRNAs as biomarkers in human diseases", *IJSL.*, , vol. 4, no. 1, p.p. 190-201, Feb. 2022, doi:10.38058/ijsl.1050036.
[20] B. Smolarz, A. Durczyński, H. Romanowicz, K. Szyłło, and P. Hogendorf, "miRNAs in Cancer (Review of Literature)", *Int. J. Mol. Sci.*, vol. 23, no. 5, Mar. 2022, doi:10.3390/ijms23052805.
[21] J. O'Brien, H. Hayder, Y. Zayed, and C. Peng, "Overview of MicroRNA Biogenesis, Mechanisms of Actions, and Circulation", *Front. Endocrinol. (Lausanne)*, vol. 3, no. 9, Aug. 2018, doi:10.3389/fendo.2018.00402.
[22] G.A.D. Metcalf, "MicroRNAs: circulating biomarkers for the early detection of imperceptible cancers via biosensor and machine-learning advances," *Oncogene*, vol. 43, pp. 2135–2142, Jun. 2024, doi.org/10.1038/s41388-024-03076-3.
[23] A. Chakrabortty, D. J. Patton, B. F. Smith, and P. Agarwal, "miRNAs: Potential as Biomarkers and Therapeutic Targets for Cancer," *Genes*, vol.14, no.7, Jun. 2023, doi:10.3390/genes14071375.
[24] L. Tan, *et al*., "Circulating miRNAs as Potential Biomarkers for Celiac Disease Development", Front. Immunol., vol. 12, no. 734763, Dec. 2021, doi:10.3389/fimmu.2021.734763.
[25] F. Li, et al., "Long non-coding RNA CNALPTC1 promotes gastric cancer progression by regulating the miR-6788-5p/PAK1 pathway", *J. Gastrointest Oncol*., vol.13, no.6, p.p. 2809-2822, Dec 2022, doi: 10.21037/jgo-22-1069.
[26] C. Wang, H. Q. Fan, and Y.W. Zhang, "MiR-511-5p functions as a tumor suppressor and a predictive of prognosis in colorectal cancer by directly targeting GPR116", *Eur. Rev. Med. Pharmacol. Sci.*, vol. 23, no. 14, p.p. 6119-6130, Jul. 2019, doi: 10.26355/eurrev_201907_18425.
[27] C. Zhang, T. Yang, and H. Jiang, "miR-511 inhibits proliferation and metastasis of breast cancer cells by targeting FGF4", *J. Gene Med.*, vol. 22, no.9, Sep. 2020, doi: 10.1002/jgm.3168.
[28] W. Yong, K. Zhang, Y. Deng, W. Tang, and R. Tao, "miR-511-5p Suppresses Cell Migration, Invasion and Epithelial-Mesenchymal Transition Through Targeting PAK2 in Gastric Cancer", *Biochem. Genet.*, vol. 60, no.3, p.p. 899-913, Jun. 2022, doi:10.1007/s10528-021-10126-y.
[29] Z. Fang, *et al.*, "Regulation of TRIM24 by miR-511 modulates cell proliferation in gastric cancer", *J. Exp. Clin. Cancer Res.*, vol. 36, no. 1, Jan. 2017, doi:10.1186/s13046-017-0489-1.
[30] H. H. Zhang, *et al*., "miR-511 induces the apoptosis of radioresistant lung adenocarcinoma cells by triggering BAX", *Oncol. Rep*., vol. 31, no.3, p.p. 1473–1479, Mar. 2014, doi:10.3892/or.2014.2973.