IJHSR
ijhighschoolresearch.org
RESEARCH ARTICLE

Unraveling a CircRNA-MiRNA-mRNA Axis: A Potential Therapeutic Target for Non-small Cell Lung Cancer

Sahil Patel

St. Anthony's High School, 275 Wolf Hill Road, Huntington Station, New York, 11747, USA; sahil.p.2008@gmail.com
Mentor: Madhav Subramanian

ABSTRACT: Non-small cell lung cancer (NSCLC) is the leading cause of cancer-related mortality, primarily attributed to NSCLC's early detection rate of approximately 19%. The circRNA-miRNA-mRNA pathway contributes to NSCLC and could be an optimal biomarker/target; its identification can significantly increase survival. From CircNET, CircFNDC3B was found to affect NSCLC development, miR-885-3p had a confidence rating of 3/3, and its target genes showed a significant impact on lung cancer patients' survival (p=0.011) on GEPIA 2 using LUAD and LUSC data sets. The TCGA lung data set was then downloaded into R, and using GSVA, the expression of miR-885-3p and specific genes was calculated. Using GEPIA 2, target gene MCM5 negatively correlated with miRNA expression in lung cancer patients (r = -0.281) and impacted NSCLC survival (p=0.0018). A gene ontology analysis on circFNDC3B and literature was then used to find GSEA gene sets. Enrichment scores were found using GSVA in R Studio with the TCGA lung data set. MCM5 expression was divided into high and low and was correlated with the GSEA gene set enrichment scores to identify differential expression within lung cancer. CircFNDC3B-miR-885-3p-MCM5 axis significantly regulates Ras protein signal transduction and IGF transport and uptake by IGFBP in NSCLC.
KEYWORDS: Biomedical and Health Sciences, Genetics and Molecular Biology of Disease, Non-Small Cell Lung Cancer, Circular RNA, MicroRNA, Messenger RNA.

Introduction

Non-small cell Lung Cancer:

Lung cancer is the second most commonly diagnosed cancer and the leading cause of cancer-related death. Non-small cell lung cancer (NSCLC) is the most common form of lung cancer, being about 85% of all diagnosed types, as stated by Huang *et al.*[1-3] Early diagnosis of NSCLC is vital for survival, drastically changing survival rates from approximately 75% in early diagnosed (stages zero, one, and two) NSCLC to less than 10% (stages three and four). However, early detection rates (detection of NSCLC in stages zero, one, and two) are drastically low at approximately 19%.[4,5] Recently, early detection has been improved by improvements in molecular-driven detection, primarily through regulatory RNAs acting as biomarkers.[6]

Current Biomarkers and Potential NSCLC Biomarkers:

Biomarkers play crucial roles in diagnosis, prognosis, and selection of treatment options. Identifying biomarkers can lead to an increase in early detection by faster diagnosis, and is also necessary for identifying the effectiveness of treatments and specific subtypes of cancers.[7] Circulating biomarkers are easy to measure and primarily consist of non-coding RNA (ncRNA), which are very sensitive and highly stable.[7] There are two main subtypes of ncRNA: long ncRNA and short ncRNA. MicroRNA (miRNA) is a form of short ncRNA and has a significant role in gene regulation through messenger RNA (mRNA) degradation.[8] Numerous research studies have found deregulated miRNAs within the plasma and serum of cancer patients, and over 2500 miRNAs have been identified as being involved in aspects of cancer development within humans.[9] miRNAs often show significant diagnostic value due to their strong specificity, high stability, and availability. This is due to miRNAs' presence in body fluids and regulation of many stages of cancer development.[10] Two miRNA signatures with the potential to be biomarkers for early diagnosis of lung cancer, miR-33a-5p and miR-128-3p, were identified. They found miR-33a-5p (AUC = 0.8644, 95% confidence interval (CI) = 0.8016 to 0.9271, sensitivity = 84.62% and specificity = 76.92%) and miR-128-3p (AUC = 0.9391, 95% CI = 0.9199 to 0.9835, sensitivity = 92.31% and specificity = 83.08%) to have a higher diagnostic value than traditional tumor markers. Traditional tumor markers within lung cancer showed lower diagnostic values such as cytokeratin-19-fragment (CYFR21-1) (AUC = 0.5856, 95% CI = 0.4387 to 0.7324, sensitivity = 63.33% and specificity = 63.33%), neuron-specific enolase (NSE) (AUC = 0.6189, 95% CI = 0.4748 to 0.763, sensitivity = 73.33% and specificity = 56.67%), and cancer antigen 72-4 (CA72-4) (AUC = 0.5206 95% CI = 0.3684 to 0.6727, sensitivity = 86.67% and specificity = 36.67%).[11] This suggests that miRNAs could potentially be used as biomarkers in lung cancer due to their greater ability to identify lung cancer than previously used biomarkers.[8] Further investigation found dysregulated miRNAs within NSCLC. Three miRNAs were specifically deregulated only in Lung Adenocarcinoma (LUAD); miR-6785-3p was upregulated, and miR-101-3p and miR-139-5p were downregulated. Five miRNAs were specifically deregulated in Lung Squamous Cell Carcinoma (LUSC); miRNA-21-3p and miRNA-650 were upregulated, and miRNA-95-5p, miRNA-4639-3p, and miR-744-3p were downregulated. Five miRNAs were commonly deregulated in both LUAD and LUSC; miR-7-5p was up-

DOI: 10.36838/IJHSR810.1

regulated, miR-140-3p, miR-144-3p, and miR-195-5p were downregulated, and miR-375 was seen to be upregulated in LUAD and downregulated in LUSC samples. This shows how different miRNAs can impact and affect specific tissues and cancer types, making them ideal for use as a cancer biomarker.[12] Circular RNA (circRNA) has been associated with interacting with miRNA in NSCLC through miRNA sponging. CircRNA has the potential to be an optimal biomarker/target due to its high stability and tissue specificity, showing potential for further investigation.[13,14]

CircRNAs' effect on NSCLC:

CircRNAs' properties of high resistance to exonuclease degradation lead to much greater stability compared to linear RNAs. CircRNAs are also highly specific in tissues and diseases, which can lead to increased concentrations in specific areas. CircRNA is often distinctly expressed in cancerous and noncancerous cells.[15] These traits lead circRNA to have the potential to be an optimal biomarker/target.

CircRNAs' functions within the cell are microRNA sponging, transcriptional regulation, acting as a protein template, and a few directly translating proteins.[16,17] CircRNAs' primary function is miRNA sponging; circRNA sponges the miRNA by having complementary binding sites. This then forms complexes through hybridization, forming a double-stranded molecule of two complementary single-stranded DNA/RNA molecules. Through circRNA sponging miRNA, the circRNA takes the miRNA away from its primary function of mRNA regulation. This process is often dysregulated due to the deregulation of circRNA, often by altered RNA back-splicing (upstream exon covalently links to the acceptor sequence of a downstream exon) and histone modifications. When this dysregulation occurs, the miRNA is up/downregulated, leading to down/upregulated mRNA; this affects gene expression regulation and causes adverse interactions with RNA-binding proteins (RBP), transcriptional factors, and other proteins, impacting other oncogenic pathways.[18] The dysregulation of the circRNA-miRNA-mRNA has been associated with cancer development through influencing proliferation, tumorigenesis, epithelial-to-mesenchymal transition (EMT), and metastasis.[19-21] Through identifying the circRNA-miRNA-mRNA pathway, knowledge is gained on the effects of particular molecules and the ability to use those molecules for early diagnosis and specific treatments. Due to the limited studies about the circRNA-miRNA-mRNA pathway and its known impacts on oncogenic traits, further investigation is needed.

CircRNA-miRNA-mRNA Pathways in NSCLC:

The dysregulation of the circRNA-miRNA-mRNA pathway leads to various oncogenic traits.[22] Many circRNA-miRNA-mRNA pathways have been found to influence NSCLC. The circ_HIPK3/miR149/Mammalian forkhead box transcription factor (FOXM1) pathway has been found to influence proliferation, migration, invasion, and apoptosis within lung cancer cells. Circ_HIPK3 was found to sponge miR-149, and miR-149 interacts with FOXM1, binding to the 3'UTR site. FOXM1 is a master regulator of tumor metastasis, giving miR-149 a tumor-suppressive role in lung cancer. Circ_HIPK3 has also been associated with regulating miR-124, which impacts target genes SphK1, STAT3, and CDK4 in other cancers.[19,20] Both the circP4HB/miR-113a-5p/vimentin axis and the circPTPRA/miR-96-5p/RASSF8 axis have been shown to lead to an increase in EMT and metastasis within NSCLC. Circ_0067934 regulates the miR-1182/KLF8 axis, influences the Wnt/beta-catenin pathway, and is associated with NSCLC development. Circ_000567/miR-421/TMEM100 axis is involved in the migration and invasion of lung adenocarcinoma.[21] F-circEA1 regulates its parental gene EML4-ALK, which is a fusion protein that impacts downstream signaling pathways such as RAS-RAF-MEK-ERK, P13K-AKT-mTOR, and JAK3-STAT3; the F-circEA1/EML4-ALK axis is involved in proliferation, migration, invasion, and tumor progression.[22] Circ-Foxo3 regulates its parent gene, Foxo3. The Circ-Foxo3/miR-155/Foxo3 axis has been found to influence cell proliferation, migration, and invasion with NSCLC cells.[23]

Circ_0020714/miR-30a-5p/SOX4 has been seen to be involved in immune evasion and resistance to antiPD-1 and showed potential use as an immunotherapeutic target in NSCLC.[24] The dysregulation of the circRNA-miRNA-mRNA pathway often affects multiple oncogenic pathways and traits. Many signaling protein pathways influence NSCLC development, but often, how these pathways are deregulated, such as by circRNA-miRNA-mRNA pathways, and ways to inhibit these pathways are limited. Rat Sarcoma (Ras) pathway and upregulation of the P13k/AKT/mTOR pathway have been involved in proliferation within NSCLC; the inhibition of these pathways has been seen to reduce proliferation.[25,26] Histone deacetylase 1 (HDAC1) has been observed to affect EMT-dependent malignant progression of NSCLC through the influence of MCM5. Abnormal interactions of MCM5 and HDAC1 lead to the downregulation of E-cadherin and upregulation of Vimentin and MMPs; this causes EMT-related tumor metastasis, as seen in Figure 1.[27] Insulin-like growth factor binding protein 3 (IGFBP3) acts as a tumor suppressor by performing its primary function of binding to insulin-like growth factor 1 (IGF1). IGF1 binds to insulin-like growth factor receptor 1 (IGFR1), activating the P13k/AKT/mTOR pathway, stimulating cell growth, and blocking apoptosis. Increases in IGF1 concentrations are associated with increased metastasis and tumorigenesis. This means that downregulated IGFBP3 is associated with an increase in metastasis and tumorigenesis, as illustrated in Figure 1.[28]

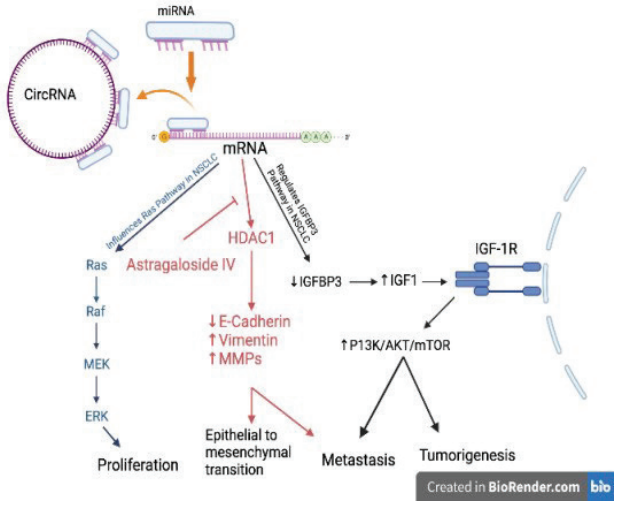

Figure 1: Dysregulated circRNA-miRNA-mRNA pathway. The effects of a dysregulated circRNA-miRNA-mRNA pathway potentially have in NSCLC development through the Ras, HDAC1, and IGFBP3 pathways. The dysregulated circRNA abnormally increases sponging of the miRNA, taking away the miRNAs' ability to suppress the mRNA, leading to a cascade of different gene pathways.

The circRNA-miRNA-mRNA pathways have been associated with regulating signaling pathways that influence oncogenic development. Investigating circRNA-miRNA-mRNA pathways can expand our understanding of cancer development and potentially lead to biomarkers or therapeutic targets in the future.

Objectives:

CircRNA has been recently associated with influencing NSCLC development primarily through the circRNA-miRNA-mRNA pathway. Diagnosis in the late stages of NSCLC has been associated with increased mortality; fully elucidating an influential circRNA-miRNA-mRNA pathway within NSCLC can give rise to biomarkers/therapeutic targets, increasing early detection rates and survival. The first objective is to identify a circRNA with high expression in NSCLC and a high-confidence miRNA target, which impacts the survival of NSCLC patients. This establishes a circRNA-miRNA pathway involved in NSCLC. The second objective is to identify the miRNA-mRNA interactions that are involved in NSCLC survival. The third and last objective is to find the mRNA's biological functions and determine the mRNA's impact on NSCLC. These objectives contribute to the elucidation of a circRNA-miRNA-mRNA pathway involved in NSCLC.

Methods

Identifying the circRNA-miRNA axis with influence on NSCLC survival:

CircNET, a database that analyzes circRNA regulatory networks in cancers, was used to find the most highly expressed circRNA (n=50) in lung cancer.[29] With this, a Gene Ontology (GO) Analysis was run to identify the circRNA's functionality in lung cancer. For each circRNA, the miRNAs it influences were analyzed to find a three-tool (PITA, miRanda, and TargetScan) confidence-level miRNA (n=14). This establishes a circRNA-miRNA regulatory network with confidence in three out of three tool types. Using the three tools, confidence miRNA, and its target genes were loaded into Gene Expression Profiling Interactive Network 2 (GEPIA 2), a database that uses The Cancer Genome Analysis (TCGA) and Genotype-Tissue Expression (GTEx) data to analyze genes in cancers.[30] Using the Lung Adenocarcinoma (LUAD) and Lung Squamous cell carcinoma (LUSC) data sets, which constitute >85% of NSCLC, a Kaplan-Meier graph (log-rank test) was created of the target genes of the miRNA.[3] This was done to determine the miRNAs' effect on NSCLC survival rates. To determine the circRNA-miRNA axis with a significant impact on NSCLC, an evaluation of the Kaplan-Meier graphs was performed to identify p-values less than the alpha value of 0.05. The selection was based on circRNA abundance, miRNA confidence, and miRNA impact/significance on NSCLC.

Identifying the target miRNA-mRNA pathway with significant influence in NSCLC:

First, CircNET was used to identify the mRNA targets of the miRNA. A detailed literature review was then conducted to identify potential mRNAs influenced by the miRNA that have an impact on cancers. GEPIA 2 was then used to identify the mRNA expression in lung cancer compared to normal lung cells; 0.5 was the Log2FC Cutoff, and the p-value was set to 0.05. Then, TCGAs' lung datasets (n = 1129) were loaded into R Studio, and the Gene Set Variation Analysis (GSVA) package was used to calculate the enrichment of gene signatures within the dataset.[31] Using GSVA, the enrichment of miRNA target genes and specific genes in NSCLC was found. After obtaining the enrichment values of the miRNA target genes and specific genes, they were loaded into an Excel sheet. The miRNA target genes enrichment values were then multiplied by -1. This is because GSVA found the enrichment of the targets of the miRNA by multiplying by -1, the proxy for miRNA expression is found; this is due to miRNA regulating its target gene, so as one increases, the other must decrease. A Pearson correlation coefficient was then found using the proxy miRNA enrichment and specific gene enrichment value in Excel. The mRNA with a significant negative correlation was identified using the correlation coefficient. Then, I ran GEPIA 2 and created a Kaplan-Meier graph to identify a specific mRNA influencing NSCLC with a p-value of <0.05. The mRNA with increased expression in lung cancer versus normal lung cells had a significant negative correlation with the given miRNA and was seen to influence NSCLC survival significantly. It was then picked for further analysis.

Identifying the functionality of the circRNA-miRNA-mRNA pathway in NSCLC:

Using the GO analysis of the circRNA function and previous literature, Gene Set Enrichment Analysis (GSEA) gene sets (n=22) were found based on functions within cancerous cells that align with the individual circRNA, miRNA, and mRNA functions previously established in the literature. Using the GSEA gene set, all genes in the sets were composed in an Excel sheet, and a GSVA was run in R Studio to find enrich-

DOI: 10.36838/IJHSR810.1

ment values of the gene sets in NSCLC.[32,33] These enrichment values and mRNA enrichment values were then placed into an Excel sheet, and the top 50% highest mRNA expression in each gene set was found. After finding the values, PRISM was used to place the enriched genes of the gene set that are in the top 50% of mRNA expression under high expression and the bottom 50% under low expression to create scatter dot plots and run t-tests to obtain p-values. These graphs were then used to assess the mRNAs' function in NSCLC. Identifying significant differential expression between low and high expression of mRNA showed that the circRNA-miRNA-mRNA axis influenced the particular function within NSCLC.

Justification of Methods:

Databases and Datasets were used due to their validity, reliability, feasibility, and sample size. The use of databases and datasets allowed for the analysis of data with a far greater sample size than in vitro/vivo experimentation would have allowed. The TCGA lung data set allowed for the use of 1129 patient data; if done in vitro/vivo, it would have taken many years for this study to occur due to permissions needed and costly tests required.

A study by Lin-lin Zhang and colleagues used a methodology similar to my research regarding the analysis of an mRNA, Minichromosome maintenance complex component 5 (MCM5), and its effects on Lung cancer.[27] Lin-Lin Zhang and colleagues aimed to identify MCM5 interaction with HDAC1 and the effects of this interaction on EMT-dependent malignant progression in Lung cancer. Using TCGA data, they compared MCM5 expression in Lung adenocarcinoma and Lung squamous cell carcinoma to normal and found the effects of MCM5 on lung adenocarcinoma survival. My study used a similar methodology, using Kaplan-Meier survival graphs and comparing cancer versus normal cells. My research differed, however, by using TCGA and GTEx data and comparing RNAs to NSCLC, increasing validity by increasing sample size and comparing RNAs to NSCLC, not just lung adenocarcinoma. My study also uses this data in a different context, using it in regard to the circRNA-miRNA-mRNA pathway rather than mRNA-mRNA interactions.

■ Results and Discussion

CircFNDC3B-miR-885-3p is highly expressed in NSCLC and is associated with decreased survival:

According to circNET, CircFNDC3B was found to be the highest expressed circRNA in lung cancer cells, as documented in Table 1. Using circNET, a Gene Ontology analysis was run on CircFNDC3B to identify its functionality in lung cancer. CircFNDC3B was seen to primarily contribute to the regulation of cell morphogenesis, Ras protein signal transduction, gland development, histone modification, and endomembrane system organization, as presented in Figure 2. CircFNDC3B was seen to show 3/3 tool confidence with miRNA-885-3p, meaning that miRNA-885-3p was predicted to be sponged by CircFNDC3B through all three tools (PITA, miRanda, TargetScan). Using GEPIA 2, miR-885-3p's target genes were then used to identify the impact the circFNDC3B-miR-885-3p axis has on NSCLC survival rate. CircFNDC3B-miR-8853p was seen to affect NSCLC survival significantly (p=0.011). Figure 3 shows that dysregulated expression of miR-885-3p led to decreased patient survival.

Table 1: Top 10 highest circRNAs expressed in lung cancer. CircFNDC3B showed the highest amount of expression compared to other circRNAs within lung cancer.

Top	CircRNA (Host Gene)	Expression
1	CircFNDC3B	2208
2	CircCDR1	2105
3	CircFAM13B	989
4	CircCAMSAP1	986
5	CircUBXN7	868
6	CircGSE1	808
7	CircHIPK3	778
8	CircXPO1	771
9	CircSCMH1	747
10	CircCDKAL1	685

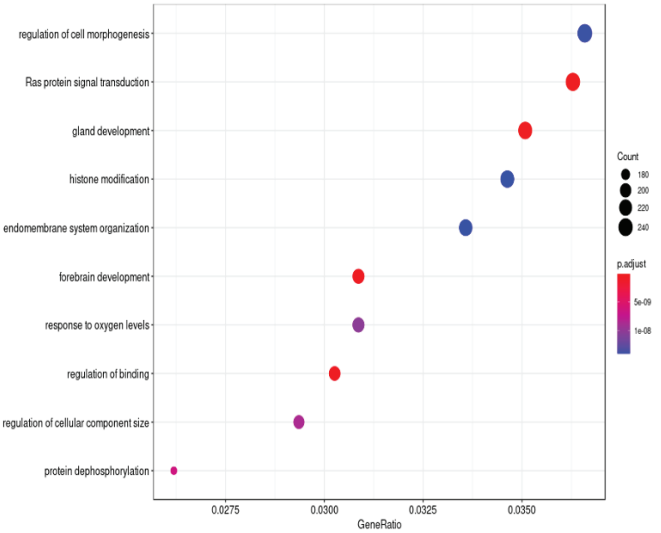

Figure 2: Functions of CircFNDC3B. Gene Ontology Analysis through CircNET to find the functionality of the circRNA-miRNA axis. The gene ratio indicates that circFNDC3B influenced enriched genes over total genes in the given pathway. As the count/size of the point increases, the amount of CircFNDC3B is seen to increase. Regulation of cell morphogenesis and Ras protein signal transduction showed the strongest influence from CircFNDC3B.

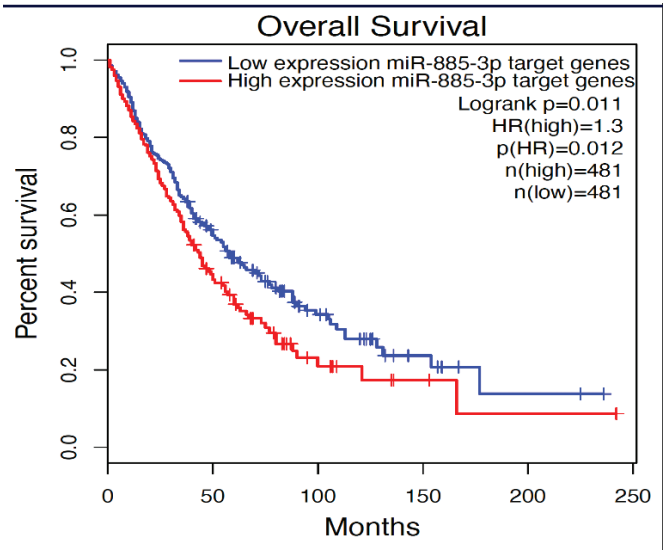

Figure 2: Overall median Kaplan-Meier graph of miR-885-3p target genes in NSCLC. Using TCGA data (LUAD and LUSC), the miR-885-3p effect in NSCLC survival was analyzed. Within LUAD and LUSC, miR-885-3p showed influence with high expression, leading to a significant increase in mortality.

CircRNA-miRNA interactions through circRNA sponging of miRNA have been seen to correlate with cancer development, from the upregulation of circRNA causing downregulation of the miRNA, leading to abnormal expression of mRNA.[34,35] CircFNDC3B has previously been correlated with being involved in cancer and disease progression.[36] CircFNDC3B is seen to exhibit characteristics of more stability, conservatism, and tissue/developmental specificity compared to linear FNDC3B.[37] In this study, CircFNDC3B was the highest expressed circRNA in lung cancer. CircFNDC3B was found to possess many functions within lung cancer, including the regulation of cell morphogenesis, Ras protein signal transduction, gland development, histone modification, and endomembrane system organization, showing the highest involvement of circFNDC3B. CircFNDC3B showed three out of three tool confidence with miR-885-3p; this correlates circFNDC3B activity with miR-885-3p, making the circFNDC3B-miR-885-3p axis. miR-885-3p has been associated with tumorigenesis within LUAD and targets the Wnt10b/βcatenin signaling pathway by regulation from circRNAs.[38] In this study, when miR-885-3p target genes were found to be increased in expression, the NSCLC survival rate decreased. This shows that circFNDC3BmiR-885-3p has a potential influence on NSCLC survival.

mRNA-MCM5 shows high expression in NSCLC and is associated with decreased survival:

MCM5 was found to be an important target of miRNA-885-3p. Previous literature showed MCM5 to have a major impact on EMT and metastasis in NSCLC, as documented in Table 2. Gepia 2, a database that uses TCGA and GTEx data, was used to find MCM5's expression in NSCLC. MCM5 was found to be significantly upregulated in LUAD and LUSC cells compared to normal lung tissue, as indicated in Figure 4. Using GSVA and TCGA lung cancer datasets (n=1129), enrichment scores were calculated for miR-885-3p target genes. To estimate miRNA activity, a proxy for miR-885-3p was generated by multiplying the target gene enrichment score by -1. Importantly, this proxy does not represent physical concentrations of miR-885-3p, but the inferred collective suppressive activity of its target gene set, reflecting functional miRNA activity at the transcriptional level. Figure 5 shows MCM5 to have a significant negative correlation with miRNA-885-3p, with a correlation of -0.281. Although statistically significant, this correlation is modest in strength, reflecting the biological complexity of MCM5 regulation and the fact that miR-885-3p contributes only partially to MCM5 expression variability. To find MCM5's impact on NSCLC, TCGA data were used to find the dysregulation of MCM5 to significantly decrease NSCLC survival, as exemplified in Figure 6 (p=0.0018).

Table 2: miRNA targets genes. The genes listed showed potential for research in NSCLC. Further investigation was conducted with each miRNA target gene; MCM5 showed the best results.

miRNA target mRNAs	NSCLC relationship
Aurora A	Associated with chemoresistance and metastasis [29]-[30]
CDK6	Overexpression is associated with tumorigenesis and EMT [31]-[32]
KRT8	Upregulation in tumors causes cell migration, cell adhesion, and drug resistance [33]
MCM5	Major role in metastasis and EMT through means of epigenetic regulation [27] and [34]
PFN1	Promotes stemness and tumor-initiating ability of cancer cells [35]-[36]

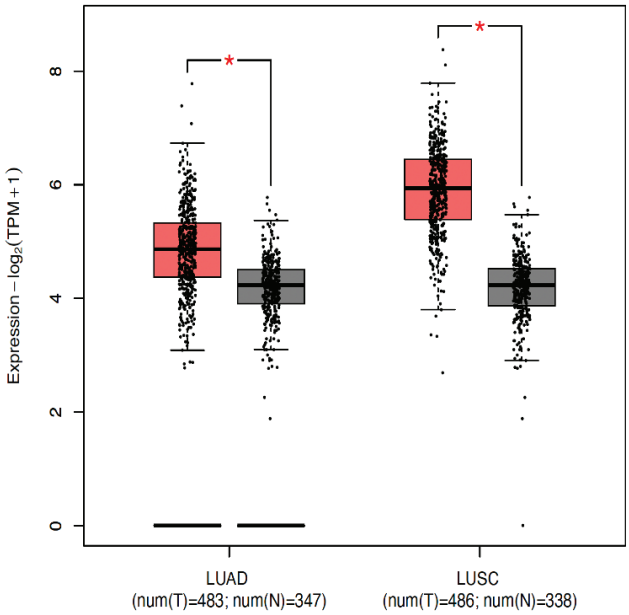

Figure 4: MCM5 expression in TCGA and GTEx NSCLC. MCM5 is significantly expressed in NSCLC compared to normal lung cells (*=<0.05). Red = Tumor lung cell; Grey = Normal lung cell.

DOI: 10.36838/IJHSR810.1

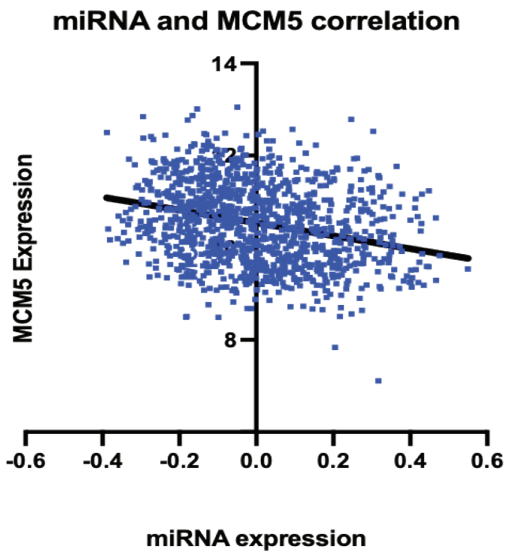

Figure 5: Correlation scatter plot. Correlation of -0.281 between miRNA target genes and MCM5 gene influence in NSCLC using TCGA data (n=1129). As miRNA-885-3p concentration increases, MCM5 concentration decreases and vice versa, indicating an inverse relationship and potentially confirming an influence on each other's expression.

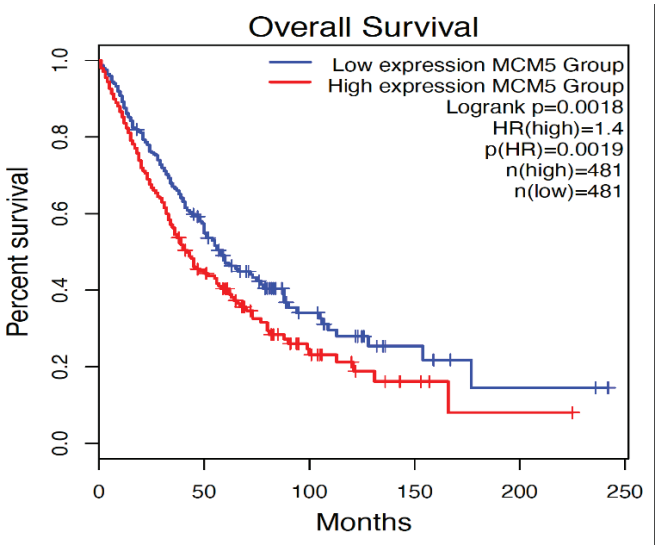

Figure 6: Overall median Kaplan-Meier graph of MCM5 on NSCLC. Using TCGA data (LUAD and LUSC), the effect of mRNA MCM5s on survival was analyzed. High expression of MCM5 was seen to significantly increase mortality rates.

A literature review of miR-885-3p's target genes was performed, and MCM5 was found to show a potential influence on cancer progression. MCM5 was seen to have a significant role in metastasis and EMT by epigenetic regulation.[27, 44] MCM5 expression in LUAD and LUSC using both TCGA and GTEx datasets showed that MCM5 is more highly expressed in LUAD and LUSC compared to normal lung tissue. MCM5 was also previously found to be more highly expressed in both LUAD and LUSC compared to normal lung tissue, although only using TCGA data.[27] MCM5 is seen to be negatively correlated with miR-885-3p; this means that as there is an increase in miR-885-3p, there is a decrease in MCM5 expression. MCM5 was then found to impact lung cancer survival when dysregulated. This means that miR-885-3p's regulation of MCM5 causes an increase in survival, but when circFNDC3B sponges miR-885-3p, it causes a decrease in expression, leading to a decrease in the regulation of MCM5 and increased expression of MCM5, causing oncogenic properties.

MCM5 shows involvement in potential oncogenic pathways:

MCM5 functionality is vital to understanding the circFNDC3B-miR-885-3p-MCM5s' impact on NSCLC. GSEA datasets were obtained using circFNDC3B's GO analysis and literature; the GSEA datasets correlate with circFNDC3B, miR-885-3p, and MCM5s' previously established functions. GSVA scores were produced to show MCM5s' high and low expression in pathways associated with circFNDC3B and NSCLC development. MCM5 showed significant differential expression between high and low expression in Ras protein signal transduction and regulation of Insulin-like growth factor (IGF) transport and uptake by Insulin-like growth factor binding proteins (IGFBP), as seen in Figures 7 and 8, respectively.

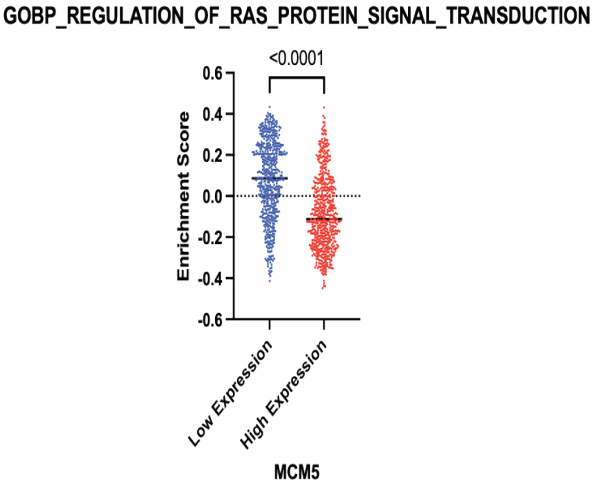

Figure 7: MCM5 expression in Ras protein signal transduction within NSCLC. MCM5 showed significant differential expression when compared to enriched genes involved in Ras protein signal transduction in NSCLC.

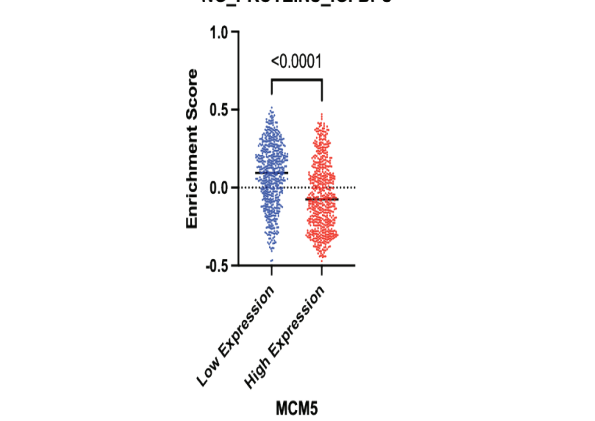

Figure 8: MCM5 expression in the regulation of insulin-like growth factor (IGF) transport and uptake by insulin-like growth factor binding proteins (IGFBP). MCM5 showed significant differential expression when compared to the enriched genes involved in IGF transport and uptake by IGFBPs in NSCLC.

The influence of MCM5 in pathways within NSCLC was then found. MCM5 was seen to significantly influence Ras protein signal transduction and regulation of IGF transport and uptake by IGFBPs. The dysregulation of Ras protein signal transduction has been associated with aggression and cellular proliferation.[47] Dysregulation of IGF transport and uptake by IGFBPs has been associated with EMT and metastasis.[48,49]

Conclusion

Overall, it was found that the CircFNDC3B/miR-885-3p/MCM5 pathway is potentially associated with NSCLC survival through significant regulation of Ras protein signal transduction and IGF transport and uptake by IGFBPs. Through the identification of the circFNDC3B/miR-885-3p/MCM5 pathway within NSCLC, knowledge has been gained of the newly emerged circRNA/miRNA/mRNA regulatory pathways, helping us understand the roles of specific RNA interactions within lung cancer.[50]

Limitations:

A potential limitation of this research is the use of datasets and databases; this is due to only using precomposed data from other scientists, which could lead to flawed results. To attempt to combat this, only highly credible and widely used datasets and databases that had the support of many studies and research from the past were chosen. The use of these datasets and databases was necessary for time, money, resources, and validity purposes. Obtaining patient lung cancer data with a large sample size would have been a complex and tedious process, taking many years to complete.

Further Research

There are many ways in which this research could be continued. Further identifying the positive correlations within the CircFNDC3B-miR-885-3p-MCM5 pathways can lead to potential biomarkers/therapeutic targets in the future to help increase NSCLC early detection rates, therefore increasing survival.

Further investigation should be done on excluded circRNAs, miRNAs, and mRNAs from NSCLC. Identifying the connections between them and the potential for pathways influencing oncogenic factors is vital for the growth of knowledge and understanding of how different pathways and RNAs interact with each other in cancers.

Further investigations should be done with the MCM5-Ras and MCM5-IGFBP interactions to determine further medical applications. This would help further correlate the interactions between the circFNDC3B-miR-885-3p-MCM5 pathway and its functions within NSCLC.

As well as the testing of potential inhibitors of this pathway for use as a biological target in the future. Astragaloside IV is a chemical compound often seen as having a protective effect on the lung; it has been found that Astragaloside IV can block interactions between HDAC1 and MCM5, inhibiting the progression of malignant lung cancer.[27]

Applications:

A potential application for this research would be the early detection of NSCLC through liquid biopsies. CircRNAs are highly abundant due to their resistance to exonuclease degradation, stability, and specificity. Due to this, they are considered novel biomarkers for liquid biopsies.[51] The presence of circRNAs in biofluids has been associated with cancer progression. This can make testing more common, helping people get diagnosed at early stages due to non-invasive and safe testing. Microarray testing is also a promising way to test for cancer development; knowledge of the circRNA axis developed in this study is beneficial, as microarray testing can only investigate known circRNAs.

Another potential application of this research is the development of therapeutic targets for NSCLC. Due to circRNAs' high resistance to exonuclease degradation and high specificity, it is seen as an optimal target.

Another application is the use of this potential axis in future research to develop further pathways and expand our knowledge of how RNAs and RNA pathways interact to cause cancer metastasis, tumorigenesis, proliferation, and more.

Acknowledgments

Thank you to Madhav Subramanian, Dr. Serena McCalla, and Mr. Paul Paino for your guidance during the development of this project.

References

1. Huang, J.; Deng, Y.; Tin, M. S.; Lok, V.; Ngai, C. H.; Zhang, L.; Lucero-Prisno, D. E.; Xu, W.; Zheng, Z.-J.; Elcarte, E.; Withers, M.; Wong, M. C. S. Distribution, Risk Factors, and Temporal Trends for Lung Cancer Incidence and Mortality. *Chest* **2022**, *161* (4), 1101–1111. https://doi.org/10.1016/j.chest.2021.12.655.
2. Sung, H.; Ferlay, J.; Siegel, R. L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA: a Cancer Journal for Clinicians* **2021**, *71* (3), 209–249.
3. Alduais, Y.; Zhang, H.; Fan, F.; Chen, J.; Chen, B. Non-Small Cell Lung Cancer (NSCLC): A Review of Risk Factors, Diagnosis, and Treatment. *Medicine* **2023**, *102* (8), e32899. https://doi.org/10.1097/md.0000000000032899.
4. Li, C.; Wang, H.; Jiang, Y.; Fu, W.; Liu, X.; Zhong, R.; Cheng, B.; Zhu, F.; Xiang, Y.; He, J.; Liang, W. Advances in Lung Cancer Screening and Early Detection. *Cancer Biology & Medicine* **2022**, *19* (5), 591–608. https://doi.org/10.20892/j.issn.2095-3941.2021.0690.
5. Zhang, C.; Ma, L.; Niu, Y.; Wang, Z.; Xu, X.; Li, Y.; Yu, Y. Circular RNA in Lung Cancer Research: Biogenesis, Functions, and Roles. *International Journal of Biological Sciences* **2020**, *16* (5), 803–814. https://doi.org/10.7150/ijbs.39212.
6. Liu, S.-Y. M.; Zheng, M.-M.; Pan, Y.; Liu, S.-Y.; Li, Y.; Wu, Y.-L. Emerging Evidence and Treatment Paradigm of Non-Small Cell Lung Cancer. *Journal of Hematology & Oncology* **2023**, *16* (1). https://doi.org/10.1186/s13045-023-01436-2.
7. Juan Carlos Restrepo; Dueñas, D.; Corredor, Z.; Liscano, Y. Advances in Genomic Data and Biomarkers: Revolutionizing NSCLC Diagnosis and Treatment. *Cancers* **2023**, *15* (13), 3474–3474. https://doi.org/10.3390/cancers15133474.

DOI: 10.36838/IJHSR810.1

8. Condrat, C. E.; Thompson, D. C.; Barbu, M. G.; Bugnar, O. L.; Boboc, A.; Cretoiu, D.; Suciu, N.; Cretoiu, S. M.; Voinea, S. C. miRNAs as Biomarkers in Disease: Latest Findings Regarding Their Role in Diagnosis and Prognosis. *Cells* **2020**, *9* (2). https://doi.org/10.3390/cells9020276.
9. He, Y.; Lin, J.; Kong, D.; Huang, M.; Xu, C.; Kim, T.-K.; Etheridge, A.; Luo, Y.; Ding, Y.; Wang, K. Current State of Circulating MicroRNAs as Cancer Biomarkers. *Clinical Chemistry* **2015**, *61* (9), 1138–1155. https://doi.org/10.1373/clinchem.2015.241190.
10. Smolarz, B.; Durczyński, A.; Romanowicz, H.; Szyłło, K.; Hogendorf, P. MiRNAs in Cancer (Review of Literature). *International Journal of Molecular Sciences* **2022**, *23* (5), 2805. https://doi.org/10.3390/ijms23052805.
11. Pan, J.; Zhou, C.; Zhao, X.; He, J.; Tian, H.; Shen, W.; Han, Y.; Chen, J.; Fang, S.; Meng, X.; Jin, X.; Gong, Z. A Two-MiRNA Signature (MiR-33a-5p and MiR-128-3p) in Whole Blood as Potential Biomarker for Early Diagnosis of Lung Cancer. *Scientific Reports* **2018**, *8* (1), 16699. https://doi.org/10.1038/s41598-01835139-3.
12. Petkova, V.; Marinova, D.; Kyurkchiyan, S.; Stancheva, G.; Mekov, E.; Kachakova-Yordanova, D.; Slavova, Y.; Kostadinov, D.; Mitev, V.; Kaneva, R. MiRNA Expression Profiling in Adenocarcinoma and Squamous Cell Lung Carcinoma Reveals Both Common and Specific Deregulated MicroRNAs. *Medicine* **2022**, *101* (33), e30027. https://doi.org/10.1097/md.0000000000030027.
13. Sufianov, A.; Begliarzade, S.; Beilerli, A.; Liang, Y.; Ilyasova, T.; Beylerli, O. Circular RNAs as Biomarkers for Lung Cancer. *Non-coding RNA Research* **2023**, *8* (1), 83–88. https://doi.org/10.1016/j.ncrna.2022.11.002.
14. Wang, F.; Yu, C.; Chen, L.; Xu, S. Landscape of Circular RNAs in Different Types of Lung Cancer and an Emerging Role in Therapeutic Resistance (Review). *International Journal of Oncology* **2022**, *62* (2). https://doi.org/10.3892/ijo.2022.5469.
15. Zhang, Z.; Yang, T.; Xiao, J. Circular RNAs: Promising Biomarkers for Human Diseases. *EBioMedicine* **2018**, *34*, 267–274. https://doi.org/10.1016/j.ebiom.2018.07.036.
16. Zhou, W.-Y.; Cai, Z.-R.; Liu, J.; Wang, D.-S.; Ju, H.-Q.; Xu, R.-H. Circular RNA: Metabolism, Functions and Interactions with Proteins. *Molecular Cancer* **2020**, *19* (1). https://doi.org/10.1186/s12943-020-01286-3.
17. Liu, Y.; Wang, L.; Liu, W. Roles of CircRNAs in the Tumorigenesis and Metastasis of HCC: A Mini Review. *Cancer Management and Research* **2022**, *Volume 14*, 1847–1856. https://doi.org/10.2147/cmar.s362594.
18. Li, J.; Zhang, Q.; Jiang, D.; Shao, J.; Li, W.; Wang, C. CircRNAs in Lung Cancer- Role and Clinical Application. *Cancer Letters* **2022**, *544*, 215810. https://doi.org/10.1016/j.canlet.2022.215810.
19. Kim, W. R.; Park, E. G.; Lee, D. H.; Lee, Y. J.; Bae, W. H.; Kim, H.-S. The Tumorigenic Role of Circular RNAMicroRNA Axis in Cancer. *International Journal of Molecular Sciences* **2023**, *24* (3), 3050. https://doi.org/10.3390/ijms24033050.
20. Lu, H.; Han, X.; Ren, J.; Ren, K.; Li, Z.; Sun, Z. Circular RNA HIPK3 Induces Cell Proliferation and Inhibits Apoptosis in Non-Small Cell Lung Cancer through Sponging MiR-149. *Cancer Biology & Therapy* **2019**, *21* (2), 113–121. https://doi.org/10.1080/15384047.2019.1669995.
21. Guo, L.; Jia, L.; Luo, L.-L.; Xu, X.; Xiang, Y.; Ren, Y.; Ren, D.; Shen, L.; Liang, T. Critical Roles of Circular RNA in Tumor Metastasis via Acting as a Sponge of MiRNA/IsomiR. *International Journal of Medical Science* **2022**, *23* (13), 7024–7024. https://doi.org/10.3390/ijms23137024.
22. Huo, Y.; Tangfeng Lv; Ye, M.; Zhu, S.; Liu, J.; Liu, H.; Song, Y. F-CircEA1 Regulates Cell Proliferation and Apoptosis through ALK Downstream Signaling Pathway in Non-Small Cell Lung Cancer. *Human Cell* **2021**, *35* (1), 260–270. https://doi.org/10.1007/s13577-021-00628-7.
23. Wei, J.; Li, M.; Xue, C.; Chen, S.; Zheng, L.; Deng, H.; Tang, F.; Li, G.; Xiong, W.; Zeng, Z.; Zhou, M. Understanding the Roles and Regulation Patterns of CircRNA on Its Host Gene in Tumorigenesis and Tumor Progression. *Experimental and Clinical Cancer Research* **2023**, *42* (1). https://doi.org/10.1186/s13046-023-02657-6.
24. Wu, J.; Zhu, M.-X.; Li, K.-S.; Peng, L.; Zhang, P.-F. Circular RNA Drives Resistance to Anti-PD-1 Immunotherapy by Regulating the MiR-30a-5p/SOX4 Axis in Non-Small Cell Lung Cancer. *Cancer Drug Resistance* **2022**, *5* (2). https://doi.org/10.20517/cdr.2021.100.
25. Fu, N.; Xi, R.; Shi, X.; Li, R.; Zhang, Z.; Li, L.; Zhang, G.; Wang, F. Hexachlorophene, a Selective SHP2 Inhibitor, Suppresses Proliferation and Metastasis of KRAS-Mutant NSCLC Cells by Inhibiting RAS/MEK/ERK and PI3K/AKT Signaling Pathways. *Toxicology and Applied Pharmacology* **2022**, *441*, 115988. https://doi.org/10.1016/j.taap.2022.115988.
26. Tan, A. C. Targeting the PI3K/Akt/mTOR Pathway in Non-Small Cell Lung Cancer (NSCLC). *Thoracic Cancer* **2020**, *11* (3), 511–518. https://doi.org/10.1111/1759-7714.13328.
27. Zhang, L.-L.; Li, Q.; Zhong, D.; Zhang, W.; Sun, X.; Zhu, Y. MCM5 Aggravates the HDAC1-Mediated Malignant Progression of Lung Cancer. *Frontiers in Cell and Developmental Biology* **2021**, *9*. https://doi.org/10.3389/fcell.2021.669132.
28. Kuhn, H.; Frille, A.; Petersen, M. A.; Oberhuber-Kurth, J.; Hofmann, L.; Gläser, A.; Taubenheim, S.; Klagges, S.; Kraemer, S.; Broschewitz, J.; von Laffert, M.; Wirtz, H. IGFBP3 Inhibits Tumor Growth and Invasion of Lung Cancer Cells and Is Associated with Improved Survival in Lung Cancer Patients. *Translational Oncology* **2023**, *27*, 101566. https://doi.org/10.1016/j.tranon.2022.101566.
29. Chen, Y.; Yao, L.; Tang, Y.; Jhong, J.-H.; Wan, J.; Chang, J.; Cui, S.; Luo, Y.; Cai, X.; Li, W.; Chen, Q.; Huang, H.-Y.; Wang, Z.; Chen, W.; Chang, T.-H.; Wei, F.; Lee, T.-Y.; Huang, H.-D. CircNet 2.0: An Updated Database for Exploring Circular RNA Regulatory Networks in Cancers. *Nucleic Acids Research* **2021**, *50* (D1), D93–D101. https://doi.org/10.1093/nar/gkab1036.
30. Tang, Z.; Kang, B.; Li, C.; Chen, T.; Zhang, Z. GEPIA2: An Enhanced Web Server for Large-Scale Expression Profiling and Interactive Analysis. *Nucleic Acids Research* **2019**, *47* (W1), W556–W560. https://doi.org/10.1093/nar/gkz430.
31. Goldman, M. J.; Craft, B.; Hastie, M.; Repečka, K.; McDade, F.; Kamath, A.; Banerjee, A.; Luo, Y.; Rogers, D.; Brooks, A. N.; Zhu, J.; Haussler, D. Visualizing and Interpreting Cancer Genomics Data via the Xena Platform. *Nature Biotechnology* **2020**, *38* (6), 675–678. https://doi.org/10.1038/s41587-020-0546-8.
32. Mootha, V. K.; Lindgren, C. M.; Eriksson, K.-F.; Subramanian, A.; Sihag, S.; Lehar, J.; Puigserver, P.; Carlsson, E.; Ridderstråle, M.; Laurila, E.; Houstis, N.; Daly, M. J.; Patterson, N.; Mesirov, J. P.; Golub, T. R.; Tamayo, P.; Spiegelman, B.; Lander, E. S.; Hirschhorn, J. N.; Altshuler, D. PGC-1α-Responsive Genes Involved in Oxidative Phosphorylation Are Coordinately Downregulated in Human Diabetes. *Nature Genetics* **2003**, *34*(3), 267–273. https://doi.org/10.1038/ng1180.
33. Subramanian, A.; Tamayo, P.; Mootha, V. K.; Mukherjee, S.; Ebert, B. L.; Gillette, M. A.; Paulovich, A.; Pomeroy, S. L.; Golub, T. R.; Lander, E. S.; Mesirov, J. P. Gene Set Enrichment Analysis: A Knowledge-Based Approach for Interpreting Genome-Wide Expression Profiles. *Proceedings of the National Academy of Sciences* **2005**, *102* (43), 15545–15550. https://doi.org/10.1073/pnas.0506580102.

34. Ma, Y.; Zou, H. Identification of the CircRNA-MiRNA-MRNA Prognostic Regulatory Network in Lung Adenocarcinoma. *Genes* **2022**, *13* (5), 885. https://doi.org/10.3390/genes13050885.
35. Huang, J.; Yu, S.; Ding, L.; Ma, L.; Chen, H.; Zhou, H.; Zou, Y.; Yu, M.; Lin, J.; Cui, Q. The Dual Role of Circular RNAs as MiRNA Sponges in Breast Cancer and Colon Cancer. *Biomedicines* **2021**, *9* (11), 1590. https://doi.org/10.3390/biomedicines9111590.
36. Sun, K.; Yao, H.; Zhang, P.; Sun, Y.; Ma, J.; Xia, Q. Emerging Landscape of CircFNDC3B and Its Role in Human Malignancies. *Frontiers in Oncology* **2023**, *13*. https://doi.org/10.3389/fonc.2023.1097956.
37. Kristensen, L. S.; Jakobsen, T.; Hager, H.; Kjems, J. The Emerging Roles of CircRNAs in Cancer and Oncology. Nature Reviews. *Clinical Oncology* **2022**, *19* (3), 188–206. https://doi.org/10.1038/s41571-021-00585-y.
38. Yang, Y.; Fan, X.; Nie, Y.; Liu, D.; Zhu, D.; Wu, K.; Zhang, Y.; Li, W.; Tian, X.; Wang, H.; Fan, Y. CircTUBGCP3 Facilitates the Tumorigenesis of Lung Adenocarcinoma by Sponging MiR-885-3p. *Cancer Cell International* **2021**, *21* (1). https://doi.org/10.1186/s12935-021-02356-2.
39. Cao, J.; Geng, J.; Chu, X.; Wang, R.; Huang, G.; Chen, L. MiRNA-885-3p Inhibits Docetaxel Chemoresistance in Lung Adenocarcinoma by Downregulating AuroraA. *Oncology Reports* **2018**, *41* (2). https://doi.org/10.3892/or.2018.6858.
40. Qian, F.; Lin, Y.; Zhang, M.; Guo, J.; Liu, Y. Circ_0061265 Competitively Binds to MicroRNA-885-3p to Promote the Development of Gastric Cancer by Upregulating AURKA Expression. *Cancer Cell International* **2022**, *22* (1). https://doi.org/10.1186/s12935-022-02646-3.
41. Gong, W.; Wang, L.; Zheng, Z.; Chen, W.; Du, P.; Zhao, H. Cyclin-Dependent Kinase 6 (CDK6) Is a Candidate Diagnostic Biomarker for Early Non-Small Cell Lung Cancer. *Translational Cancer Research* **2020**, *9* (1), 95–103. https://doi.org/10.21037/tcr.2019.11.21.
42. Lin, Z.; Zhou, Z.; Guo, H.; He, Y.; Pang, X.; Zhang, X.; Liu, Y.; Ao, X.; Li, P.; Wang, J. Long Noncoding RNA Gastric Cancer-Related LncRNA1 Mediates Gastric Malignancy through MiRNA-885-3p and Cyclin-Dependent Kinase 4. *Cell Death & Disease* **2018**, *9* (6). https://doi.org/10.1038/s41419-018-0643-5.
43. Xie, L.; Dang, Y.; Guo, J.; Sun, X.; Xie, T.; Zhang, L.; Yan, Z.; Amin, H.; Guo, X. High KRT8 Expression Independently Predicts Poor Prognosis for Lung Adenocarcinoma Patients. *Genes* **2019**, *10* (1), 36. https://doi.org/10.3390/genes10010036.
44. Mao, J.; Shen, J.; Lu, X.; Cai, Y.; Tao, R.; Deng, Y.; Zhang, Y.; Wu, Y.; Chen, W. MCM5 Is an Oncogene of Colon Adenocarcinoma and Promotes Progression through Cell Cycle Control. *Acta Histochemica* **2023**, *125* (6), 152072. https://doi.org/10.1016/j.acthis.2023.152072.
45. Jiang, C.; Ding, Z.; Joy, M.; Chakraborty, S.; Kim, S. H.; Bottcher, R.; Condeelis, J.; Singh, S.; Roy, P. A Balanced Level of Profilin-1 Promotes Stemness and Tumor-Initiating Potential of Breast Cancer Cells. *Cell Cycle* **2017**, *16* (24), 2366–2373. https://doi.org/10.1080/15384101.2017.1346759.
46. Wang, Y.; Wang, Y.; Wan, R.; Hu, C.; Lu, Y. Profilin 1 Protein and Its Implications for Cancers. *ONCOLOGY* **2021**, *35* (3507), 402–409. https://doi.org/10.46883/onc.2021.3507.0402.
47. Alam, M.; Hasan, G. M.; Eldin, S. M.; Adnan, M.; Riaz, M. B.; Islam, A.; Khan, I.; Hassan, Md. I. Investigating Regulated Signaling Pathways in Therapeutic Targeting of Non-Small Cell Lung Carcinoma. *Biomedicine & Pharmacotherapy* **2023**, *161*, 114452. https://doi.org/10.1016/j.biopha.2023.114452.
48. Kerr, A.; Baxter, R. C. Noncoding RNA Actions through IGFs and IGF Binding Proteins in Cancer. *Oncogene* **2022**, *41* (25), 3385–3393. https://doi.org/10.1038/s41388-022-02353-3.
49. Yang, X.; Bai, Q.; Chen, W.; Liang, J.; Wang, F.; Gu, W.; Liu, L.; Li, Q.; Chen, Z.; Zhou, A.; Long, J.; Tian, H.; Wu, J.; Ding, X.; Zhou, N.; Li, M.; Yang, Y.; Cai, J. M6A-Dependent Modulation via IGF2BP3/MCM5/Notch Axis Promotes Partial EMT and LUAD Metastasis. *Advanced Science* **2023**, *10* (25). https://doi.org/10.1002/advs.202206744.
50. Jia, S.; Yu, L.; Wang, L.; Peng, L. The Functional Significance of CircRNA/MiRNA/MRNA Interactions as a Regulatory Network in Lung Cancer Biology. *The International Journal of Biochemistry & Cell Biology* **2024**, *169*, 106548–106548. https://doi.org/10.1016/j.biocel.2024.106548.
51. Zhang, Y.; Wang, Y.; Su, X.; Wang, P.; Lin, W. The Value of Circulating Circular RNA in Cancer Diagnosis, Monitoring, Prognosis, and Guiding Treatment. *Frontiers in Oncology* **2021**, *11*. https://doi.org/10.3389/fonc.2021.736546.

Author

Sahil Patel, a senior at St. Anthony's High School, aspires to work in biomedical computation, specifically computational biology and AI/ML for medicine, as well as global health policy and biotech entrepreneurship. He is an incoming biomedical engineering student and Stamps Scholar at Georgia Institute of Technology. An Eagle Scout, exchange ambulance employee, varsity tennis captain, and president and co-founder of Listen to Lead, a 501(c)(3) nonprofit, he has interned at Cold Spring Harbor Laboratory in the Hammell Lab and at Hofstra University in the Lynch Lab, further advancing his interests in computational biology and biomedical research.