

Investigating the Role of Differentially Expressed Genes and Sex-Specific Genes in Lung Adenocarcinoma: A Transcriptomic Analysis

By Harriet Lai

Author Biography

Harriet Lai is a senior in high school located in Irvine, California, and is interested in biotechnology and biomedical engineering, especially the field of bioinformatics. Harriet hopes that she can continue pursuing this type of research beyond high school. Besides research, she can be found playing basketball with her family and watching her favorite NBA team, the Golden State Warriors.

Abstract

Lung cancer is a life-threatening disease that affects hundreds of thousands of patients every year. While conventional medicine offers effective treatments for lung cancer, the use of targeted therapy can offer patients a more favorable treatment by being as effective and by minimizing the side effects the patient experiences. The purpose of this research is to identify any differentially expressed or sex-specific genes in lung adenocarcinoma (LUAD) that can act as potential targets for treatment. Using transcriptomic profiles of LUAD tissues from The Cancer Genome Atlas, these profiles are analyzed using the TCGABiolinks library on R. The differentially expressed genes are then used to generate gene ontology, create protein association networks, and identify survival plots of genes using Metascape, STRING, and the Human Protein Atlas respectively. Similar procedures were used for the sex-specific analysis. The gene ontology analysis showed that the TFAP2 family of transcription factors and the PID HNF3B pathway likely promote tumorigenesis in LUAD. Moreover, AGER and SFTPC, genes involved in lung development, were underexpressed in all LUAD samples. Patients with a higher expression of these genes had a higher survival rate than others and can act as potential prognostic markers. In sex-specific genes, it was found that members of the MAGE-A family, a testis antigen family, were highly up-regulated in both males and females. The dysregulation of these genes in females might have a specific role in LUAD progression. These findings can be investigated further to develop better therapies to treat the disease.

Keywords: lung cancer, lung adenocarcinoma, differentially expressed genes, sex-specific genes, gene ontology, tumorigenesis, prognostic genes, targeted therapy

Introduction

In 2023 alone, an estimated 238,000 new cases of lung cancer and 127,000 lung cancer-related deaths were recorded in the United States (American Cancer Society, n.d.). Lung cancer is the leading cause of cancer-related deaths in both men and women. The 5-year survival rate can dip as low as 9% if a patient reaches Stage IV lung cancer (VeryWell Health, n.d.). Patients may exhibit many symptoms such as shortness of breath, coughing of blood, chest pain, and loss of weight (Mayo Clinic, 2021). There are many risk factors for lung cancer such as exposure to drugs with side effects, radiation, air pollution, and family history, but the majority of people with lung cancer have a history of smoking. As of 2020, it was found that over 87% of lung cancer patients were either current smokers or former smokers (Siegel et al., 2020). However, it is important to note that with the increasing awareness of the dangers of smoking, incidence rates for lung cancer have been falling every year since 2006 (Cancer.net, 2019).

Lung cancer isn't a singular cancer, but rather a category of cancers. 85% of all lung cancer cases fall under non-small-cell lung cancer (NSCLC). Within NSCLCs, there are two types: lung adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC). There are several differences between LUAD and LUSC, including tumor location, pathological stage, histologic grade, and treatment (Wang et al., 2019). In LUAD, the cancer begins in the cells that line the alveoli while LUSC starts in the lining of the bronchi. LUAD grows slowly while LUSC spreads quickly to other parts of the body, making it harder to treat (Yale Medicine, n.d.). As the gene profiles of these different cancer types differ, it is important to study them separately. Since LUAD is the more common of the two, my research looks only at the genetic profiles of LUAD tissue.

The primary mechanism underlying the impact lung cancer has on the body is its role in inducing changes in gene expression. Studying this further could unveil significant insights into how the disease alters the body and how the body responds to it. However, it is not as straightforward as simply drawing conclusions after comparing the gene expression of tumor and healthy tissue samples. There are other factors to consider. Past research has shown that the genetic disparity between patients determined by race and sex can influence disease progression

and patient outcomes in lung cancer. Davuluri et al. (2021) released a paper on the molecular basis of gender disparities and smoking in lung cancer patients. Using microarray and RNA-seq data, they found that the genes upregulated in female lung cancer patients played a role in the positive regulation of immune response, leukocyte and lymphocyte activation and aggregation, and cell adhesion. Interestingly, in male patients, the genes found to have a role in immune response and cell adhesion were downregulated. In another paper that used RNA-seq to compare differential gene expression in males and females by Yang et al. (2021), it was noted that the immune differences between sexes could be mainly caused by the sex chromosomes, reproductive organs, and sex hormone levels. Females were found to have a high expression of TLR7 and TLR8, part of the toll-like receptor family, crucial for the early detection of infections. Located on the short arm of the X chromosome, these genes could escape X chromosome activation, playing a role in the genes' increased expression in females. However, there is still much that is unknown about sex-related genes in lung cancer. Further research can lead to novel avenues for sex-specific cancer treatments.

Understanding which genes have major roles in LUAD can provide insight into the development of new targeted therapies. Effective treatment is crucial for preventing further progression of lung cancer. The current treatments for lung cancer include radiation therapy, which uses high-energy rays to eliminate cancer cells; chemotherapy, which involves injecting toxic chemicals to target rapidly dividing cells; and immunotherapy, which activates the immune system to target the cancer cells. They typically have unpleasant side effects and can increase the risk of developing other diseases or cancers. Therefore, personalized medicine is taking a larger role in treating diseases by specifically targeting genes that are found to be the main drivers of cancer. They can also avoid healthy cells by manipulating the expression of the gene to minimize the cause of the disease while also minimizing side effects. Therefore, it is more important than ever to study differentially expressed genes in different groups of people because these treatments would not only combat the causes and symptoms of cancer but also offer patients and doctors the time they need to decide next steps, especially for a fast-progressing and deadly disease such as lung cancer.

In this paper, the transcriptome profiles of normal lung tissue were compared to those of LUAD to identify novel differentially expressed genes that can be further analyzed in future research to determine their impact on the disease. Among the differentially expressed genes, the paper looked into potential prognostic genes that could be used for predicting the progression of the disease. The hope is that one of these genes and disease processes can be used as a target for the development of better treatments.

Materials and Methods

Acquisition of the transcriptome profile data, differential gene analysis, and visualization of volcano plots were carried out using RStudio 2023.12.0. The data is from The Cancer Genome Atlas (TCGA), a cancer genomics program between the National Cancer Institute (NCI) and the National Human Genome Research Institute, on the GDC Data Portal. Transcriptome profiling of primary tumor samples and solid tissue normal samples was accessed using GDCquery() from the TCGA-LUAD project with workflow type “STAR - Counts” and data type “Gene Expression Quantification”. The data is then downloaded using GDCdownload(). A total of 598 lung patient transcriptome profiles were downloaded, with 539 primary tumor samples and 59 solid normal tissue samples.

Analysis was carried out using the TCGAbiolinks library. After acquiring the transcriptome profiles, the data was preprocessed and normalized using TCGAanalyze_Preprocessing() and TCGAanalyze_Normalization(), respectively. Then, TCGAanalyze_Filtering() was used to apply a quantile filter with a threshold mean of 0.25. The list of differentially expressed genes was generated using TCGAanalyze_DEA() with false discovery rate (FDR) < 0.01 and log fold change (logFC) > 0.01. To visualize the differentially expressed genes, a volcano plot was generated using TCGAvisualize_volcano() with the x-axis as the logFC and the y-axis as the $-\log_{10}(\text{p-value})$. The x-axis cutoff was 0 and the y-axis cutoff was 0.01.

To identify the biological processes and cellular functions associated with the differentially expressed genes, gene ontology terms were generated using the top 100 up- and down-regulated genes,

based on log fold change with p-value < 0.05 using the platform, Metascape (<https://metascape.org>). Then, additional analysis was performed on the same top 100 up- and down-regulated genes using STRING (string-db.org), to identify functional protein association networks.

Using the Human Protein Atlas (HPA), which also generates its information from LUAD data in TCGA, the survival plots of the top 100 up- and down-regulated genes were identified. Differentially expressed genes identified in this analysis were compared against known LUAD prognostic genes using HPA to determine the prognostic value of the identified genes.

The same libraries and functions were used to generate the differentially expressed genes of females and males, but instead of one single analysis, there were two separate analyses. The female analysis used 50 female primary tumor samples and the 59 solid tissue normal samples. Similarly, the male analysis used 50 male primary tumor samples and the 59 solid tissue normal samples. A Venn diagram using the edgeR library was generated to compare the DEGs of females and males with lung adenocarcinoma.

Results and Discussion

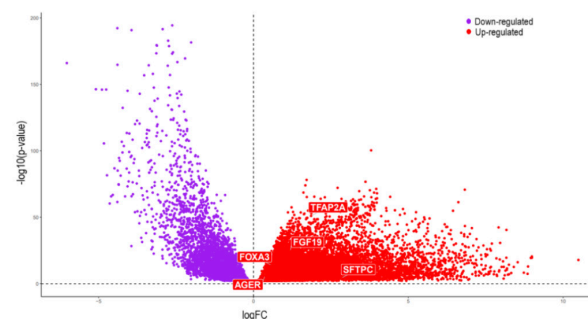


Fig 1. All the differentially expressed genes were graphed on this volcano plot. The x-axis is the log fold change (logFC) and the y-axis is the $-\log_{10}(\text{p-value})$.

Figure 1 shows that the down-regulated genes have a much wider range of significance compared to the up-regulated genes, with many down-regulated genes whose $-\log_{10}(\text{p-value}) > 50$ compared to the up-regulated genes, which are mostly less significant.

However, the up-regulated genes have a wider range of logFC with several genes beyond a logFC of 5. This could be because many normal functions are consistently down-regulated in cancer, so it is more significant. However, the up-regulated genes have a wider range of expression but are also less consistent and therefore, less significant in cancer cells. After a cutoff of $|\log FC| > 1.5$ and $p\text{-value} < 0.05$, the differentially expressed gene list contained 6592 genes, including 5438 up-regulated genes and 1154 down-regulated genes. Among the list of genes are up-regulated genes like FGF19 (logFC = 8.99, $p\text{-value} \leq 0.01$) and down-regulated genes such as AGER (logFC = -4.90, $p\text{-value} \leq 0.01$) and SFTPC (logFC = -4.60, $p\text{-value} \leq 0.01$). These genes highlighted will be discussed further.

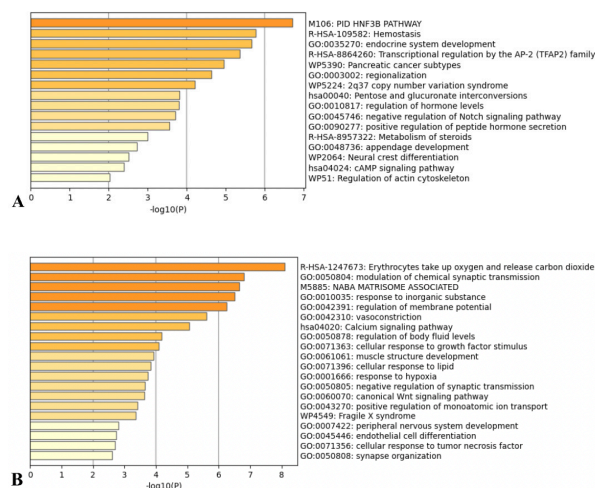


Fig 2. A) The gene ontology of the 100 up-regulated genes. B) The gene ontology of the 100 down-regulated genes.

There were a few notable processes and pathways in the results that could explain its impact or response to the disease. In Figure 2A, the PID HNF3B pathway had the highest significance in the gene ontology of the top up-regulated genes as the protein network for the FOXA2 and FOXA3 (logFC = 3.36, $p\text{-value} \leq 0.01$) transcription factors, which increase fatty acid metabolism (Gene Set Enrichment Analysis, n.d.). Evidence shows that fatty acids have a significant role in cancer cell survival and progression by providing energy and resources to cancer cells. Unusual fatty acid metabolism levels have been noted in many cancers and are found to promote metastasis

(Du et al., 2023; Chen & Huang, 2019). Regulation by the TFAP2 family of transcription factors also promotes tumor growth. The overexpression of the genes in the pathway results in less binding of TFAP2A to DNA. This leads to increased PPAR β/δ expression, which is normally repressed by TFAP2A (Jin et al., 2023). PPAR β/δ is found to promote tumor growth in lung cancer (Sun et al., 2009). Interestingly, TFAP2A (logFC = 3.79, $p\text{-value} \leq 0.01$) is overexpressed, which might indicate that other unidentified genes have a role in this process. However, not all processes and pathways identified promote cancer. Others are a response to the alterations in the body due to the cancer. The hemostasis pathway also contains many upregulated genes and functions to stop blood loss when blood leaves the blood vessels. Coughing up blood is one of the first signs of lung cancer and 20% of all patients develop this symptom (VeryWell Health, 2022). The endocrine system development indicates a change in hormone expression as a response to the cancer. Paraneoplastic syndromes like humoral hypercalcemia of malignancy, ectopic Cushing's syndrome, and carcinoid syndrome occur when the lung cancer cells produce hormones that affect the hormone balance. These pathways are found to be poor prognostic markers, but signs of these syndromes can be early signs of lung cancer and can lead to early diagnosis (Cancer Research UK, 2014; Efthymiou et al., 2018).

Among the gene ontology for down-regulated genes in Figure 2B, most processes were likely a response to the cancer. The most significant pathway describes the transfer of oxygen and carbon dioxide, facilitated by red blood cells (Reactome, n.d.). The cellular response to tumor necrosis factor (TNF) is also down-regulated. TNF, a cytokine, regulates immune cells and induces apoptosis, likely driving tumorigenesis (Herreweghe et al., 2010). There are a few known genes that are found to be repeatedly up-regulated or down-regulated in lung cancer. Through survival plots, some of those genes have been identified as prognostic genes, which can be used to predict the progression of the disease and assist with selecting a treatment option.

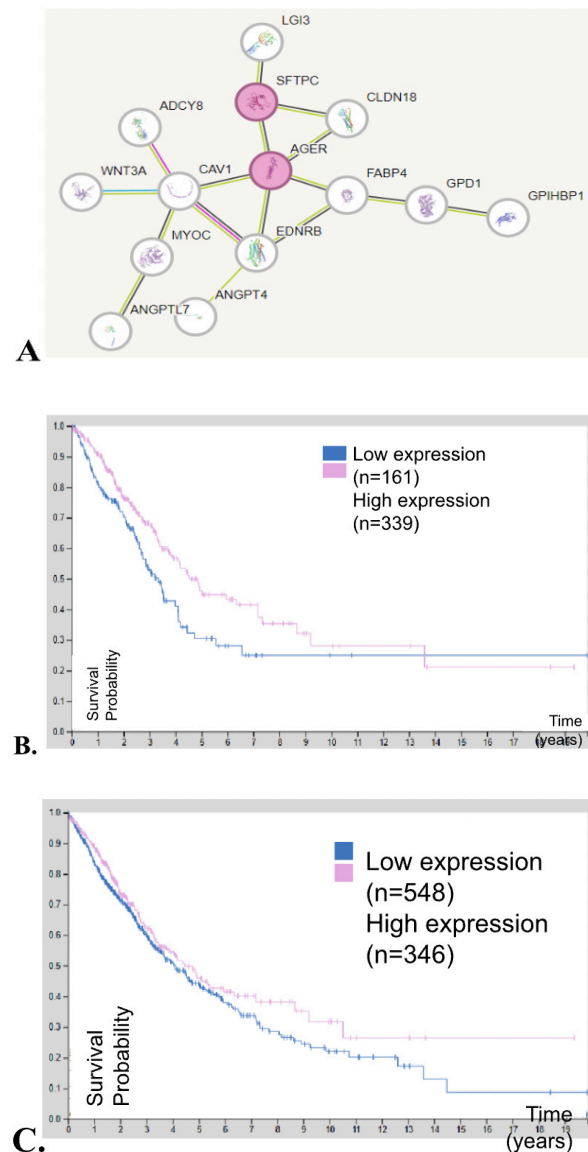


Fig 3. The survival plots were adapted from the Human Protein Atlas. A) Protein association network, generated from the top 100 down-regulated genes, highlighting the coexpression (black line) of SFTPC and AGER. B) Survival plot for primary tumor samples expressing high and low expression of SFTPC. C) Survival plot for primary tumor samples expressing high and low expression of AGER.

The protein association network in Figure 3A shows the coexpression of SFTPC and AGER, both on the list of the top 10 down-regulated genes. SFTPC (logFC = -4.596, p-value ≤ 0.01) encodes

for a surfactant that maintains pulmonary tissue stability to keep the lung from collapsing (Human Protein Atlas, n.d.). A member of the immunoglobulin superfamily of cell surface receptors, AGER (logFC = -4.89, p-value ≤ 0.01) is involved in many processes, including homeostasis, development, and inflammation (GeneCards, n.d.). The underexpression of AGER has been associated with tumor protein 53 mutation and metastasis (Yang et al., 2023). SFTPC and AGER are involved in lung development, explaining their down-regulation, and can potentially be used as markers when predicting an individual's cancer development (Mižiková & Thébaud, 2020). The survival plot for SFTPC in Figure 3B shows that those with higher expression have a higher survival probability from the beginning, with a great difference at 7 years with a 20% higher survival rate, until the 13-year mark where the trend in survival inverses. It is important to note that at this point, there is a scarce amount of data for patients that reach 9 years and beyond. Similarly, in the AGER survival plot in Figure 4C, those with higher expression have a higher survival probability.

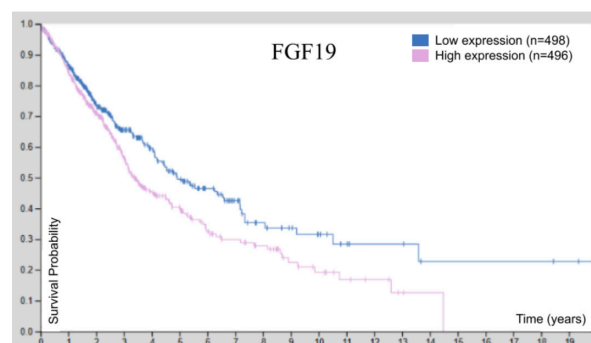


Fig 4. Survival plot for primary tumor samples expressing high and low expression of FGF19.

FGF19 (logFC = 8.99, p-value ≤ 0.01) is found to be involved in cell survival activities such as tissue repair and tumor growth. Expression of FGF19 is typically only detected in fetal brain tissue and has been found to have a role in LUSC, promoting cell growth and metastasis (GeneCards, n.d.; Zhang et al., 2022; Li et al., 2020). Unlike LUSC, much is still not known about FGF19's role in LUAD, and considering that there are many significant differences between the two types of NSCLC, FGF19 could potentially be a prognostic gene as there is a large difference in survival probability, shown in Figure 4.

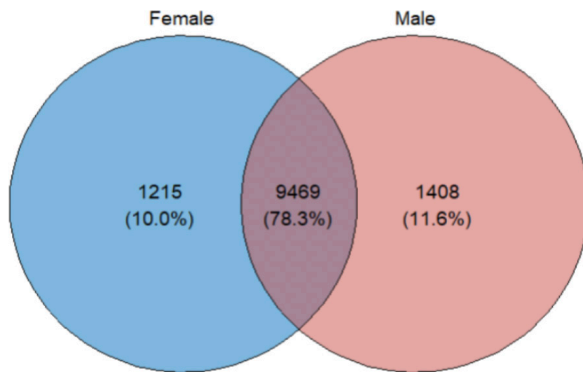


Fig 5. There are 9469 differentially expressed genes that overlap between females and males. Another 1215 genes are uniquely dysregulated for females and 1408 genes for males.

The overlap of differentially expressed genes in Figure 5 makes up a majority of the total differentially expressed genes with 9469 genes or 78.3% of all genes. This means that the other 20% of the genes are evenly split between the males and the females and are unique to the specific sex. While it is expected that the majority of the genes are shared between both sexes, there are over a thousand genes for both sexes that may have different roles that can lead to different disease progressions or other effects.

Table 2. The top 10 up- and down-regulated genes in females and males after comparing the transcriptome profiles of solid tissue normal and primary tumor tissue.

Females			Males		
Gene	logFC	p-value	Gene	logFC	p-value
MAGEA4	11.12	5.04E-14	TRIM48	10.45	1.41E-06
CGA	10.88	3.15E-16	FGF19	10.20	6.81E-21
PSG4	10.60	6.60E-14	MAGEC1	10.13	3.93E-19
PSG3	10.22	2.65E-14	MAGEA10	9.89	8.91E-17
VGLL2	10.01	2.36E-12	FGB	9.74	2.83E-30
<u>DCAF4L2</u>	9.56	4.33E-10	<u>DCAF4L2</u>	9.72	2.32E-16
FGF19	9.24	8.25E-22	MAGEA4	9.56	2.38E-14
FABP7	9.23	8.88E-23	MAGEA3	9.56	2.40E-21
MAGEC2	8.95	2.89E-18	MAGEA6	9.42	5.20E-18
NEUROD1	8.69	6.70E-15	UGT1A10	9.40	3.92E-11

Members of the MAGE-A family were found in both of the top 10 up-regulated genes for males and females (Fig. 5). This family is a testis antigen family that is typically only expressed in male germ cells. Therefore, the dysregulation of MAGE-A genes in

females should be a study point for further research on genes that are driving lung adenocarcinoma in females. Past research has found that MAGE-A has a role in the survival rate of lung adenocarcinoma patients who don't have EGFR amplification or ALK rearrangements found in many patients with NSCLC. Some treatments target EGFR and ALK, but for those who don't respond to these treatments, MAGE-A can act as an alternative target for treatment (Sang et al., 2017).

Table 3. The top 10 most significant genes in females and males after comparing the transcriptome profiles of solid tissue normal and primary tumor tissue.

Females			Males		
Gene	logFC	p-value	Gene	logFC	p-value
PYCR1	4.08	4.42E-128	CDC6	3.75	1.11E-144
TOP2A	4.40	4.84E-126	TOP2A	4.30	1.27E-139
ANLN	4.30	1.00E-116	PYCR1	4.00	9.01E-138
ASPM	4.25	4.40E-116	EXO1	3.97	2.93E-136
<u>TEDC2</u>	3.96	6.40E-114	NUF2	4.27	1.85E-133
KIF14	4.17	1.40E-113	XRCC2	3.79	1.61E-128
CENPF	3.87	5.84E-110	KIF23	2.98	1.26E-126
CDC6	3.97	8.56E-108	RS1	-5.58	4.18E-125
KIF4A	4.32	3.39E-107	NEK2	4.19	4.97E-124
STIL	2.92	3.20E-106	UBE2T	3.80	5.11E-124

Many of the most significant genes for both sexes were involved in the cell cycle which could be a factor of the fast replicating cells that are characteristic of cancers. Some of the genes encoded for microtubules or the proteins that assist microtubules. It has been found that members of the kinesin (KIF) family, including KIF4A (logFC = 4.32, p-value ≤ 0.01), are both prognostic and therapeutic targets for the cancer (Wei et al., 2023; Singharajkomron et al., 2022). While it is related to the cell cycle, it can also be a factor in metastasis. The role of microtubules in cell structure could result in abnormal cell structure that may drive the movement of cells away from their original location. Xia et al. (2024) found that MAP4's increased expression in lung adenocarcinoma regulates the epithelial-mesenchymal transition (EMT) process, which promotes the migration of diseased cells. Whether other related genes with similar functions have a similar effect is something that can be examined further.

In conclusion, this research identified biological processes and genes that play a role in LUAD, including promoting tumorigenesis and cancer cell survival. Additionally, the MAGE-A family and microtubules were found to potentially have different roles in lung progression and metastasis in males and females. The next step is to conduct research to identify which are a cause of the disease and which are an effect. Once that is identified, the hope is that there can be the development of drugs to target the causes and minimize or even eliminate the disease.

Acknowledgment

I would like to thank my mentor, Dr. James Marchant, for the constant guidance throughout my research process from offering advice while developing my research question to reviewing the paper. I'm also grateful for my TA, Garth Fisher, who was there whenever I ran across an issue and would help me resolve it. Finally, I would like to thank my parents for constantly supporting me every day.

References

- al Zeyadi, M., Dimova, I., Ranchich, V., Rukova, B., Nesheva, D., Hamude, Z., Georgiev, S., Petrov, D., & Toncheva, D. 2015. Whole genome microarray analysis in non-small cell lung cancer. *Biotechnology & Biotechnological Equipment*, 29(1), 111–118.
- American cancer society. 2023. Lung Cancer Statistics | How Common is Lung Cancer? www.cancer.org.
- American Society of Clinical Oncology. 2019. Lung Cancer - Non-Small Cell - Statistics. Cancer.net.
- Bioinformatics Pipeline: mRNA Analysis - GDC Docs. n.d. Docs.gdc.cancer.gov.
- Chen, M., & Huang, J. 2019. The expanded role of fatty acid metabolism in cancer: new aspects and targets. *Precision Clinical Medicine*, 2(3), 183–191.
- Database, G. 2017. AGER Gene - GeneCards | AGER Protein | Genecards.org.
- Database, G. 2017. FGF19 Gene - GeneCards | FGF19 Protein | FGF19 Antibody.
- Davuluri, S., Bajpai, A. K., Thirumurugan, K., & Acharya, K. K. 2021. The molecular basis of gender disparities in smoking lung cancer patients. *Life Sciences*, 267, 118927.
- Du, A., Wang, Z., Huang, T., Xue, S., Jiang, C., Qiu, G., & Yuan, K. 2023. Fatty acids in cancer: Metabolic functions and potential treatment. *MedComm – Oncology*, 2(1).
- Efthymiou, C., Spyrtos, D., & Kontakiotis, T. 2018. Endocrine paraneoplastic syndromes in lung cancer. *Hormones*, 17(3), 351–358.
- Genome Characterization Pipeline - NCI. 2022. www.cancer.gov.
- Genomic Data Commons | CRDC. n.d. Datacommons.cancer.gov.
- Jin, C., Luo, Y., Liang, Z., Li, X., Kolat, D., Zhao, L., & Xiong, W. 2023. Crucial role of the transcription factors family activator protein 2 in cancer: current clue and views. *Journal of Translational Medicine*, 21(1), 371.
- Li, F., Li, Z., Han, Q., Cheng, Y., Ji, W., Yang, Y., Lu, S., & Xia, W. 2020. Enhanced autocrine FGF19/FGFR4 signaling drives the progression of lung squamous cell carcinoma, which responds to mTOR inhibitor AZD2104. *Oncogene*, 39(17), 3507–3521.
- Lung Cancer and Hemoptysis: When to Take Action. n.d. Verywell Health.
- Mayo Clinic. 2021. Lung Cancer - Symptoms and Causes. Mayo Clinic.
- Mižíková, I., & Thébaud, B. 2021. Looking at the developing lung in single-cell resolution. *American Journal of Physiology-Lung Cellular and Molecular Physiology*, 320(5), L680–L687.
- PID_HNF3B_PATHWAY. n.d. www.gsea-msigdb.org.
- Rao, M. S., van Vleet, T. R., Ciurlionis, R., Buck, W. R., Mittelstadt, S. W., Blomme, E. A. G., & Liguori, M. J. 2019. Comparison of RNA-Seq and Microarray Gene Expression Platforms for the Toxicogenomic Evaluation of Liver From Short-Term Rat Toxicity Studies. *Frontiers in Genetics*, 9.

Reactome | Erythrocytes take up oxygen and release carbon dioxide. n.d. Reactome.org.

Sang, M., Gu, L., Yin, D., Liu, F., Lian, Y., Zhang, X., Liu, S., Huang, W., Wu, Y., & Shan, B. 2017. MAGE-A family expression is correlated with poor survival of patients with lung adenocarcinoma: a retrospective clinical study based on tissue microarray. *Journal of Clinical Pathology*, 70(6), 533–540.

SFTPC gene information - The Human Protein Atlas. n.d. www.proteinatlas.org.

Shi, H., Seegobin, K., Heng, F., Zhou, K., Chen, R., Qin, H., Manochakian, R., Zhao, Y., & Lou, Y. 2022. Genomic landscape of lung adenocarcinomas in different races. *Frontiers in Oncology*, 12, 946625.

Siegel, D. A., Fedewa, S. A., Henley, S. J., Pollack, L. A., & Jemal, A. 2021. Proportion of Never Smokers Among Men and Women With Lung Cancer in 7 US States. *JAMA Oncology*, 7(2), 302.

Singharajkomron, N., Yodsurang, V., Seephan, S., Kungsukool, S., Petchjorm, S., Maneeganjanasing, N., Promboon, W., Dangwilailuck, W., & Pongrakhananon, V. 2022. Evaluating the Expression and Prognostic Value of Genes Encoding Microtubule-Associated Proteins in Lung Cancer. *International Journal of Molecular Sciences*, 23(23), 14724.

Sun, X., Ritzenthaler, J. D., Zhong, X., Zheng, Y., Roman, J., & Han, S. 2009. Nicotine Stimulates PPAR β/δ Expression in Human Lung Carcinoma Cells through Activation of PI3K/mTOR and Suppression of AP-2 α . *Cancer Research*, 69(16), 6445–6453.

The hormone system and cancer. 2014. Cancer Research UK.

van Herreweghe, F., Festjens, N., Declercq, W., & Vandenabeele, P. 2010. Tumor necrosis factor-mediated cell death: to break or to burst, that's the question. *Cellular and Molecular Life Sciences*, 67(10), 1567–1579.

Wang, B.-Y., Huang, J.-Y., Chen, H.-C., Lin, C.-H., Lin, S.-H., Hung, W.-H., & Cheng, Y.-F. 2020. The comparison between adenocarcinoma and squamous cell carcinoma in lung cancer patients. *Journal of Cancer Research and Clinical Oncology*, 146(1), 43–52.

What Is Stage 4 Lung Cancer Life Expectancy? n.d. Verywell Health.

Wei, Y., Yang, C., Wei, J., Li, W., Qin, Y., & Liu, G. 2023. Identification and verification of microtubule associated genes in lung adenocarcinoma. *Scientific Reports*, 13(1), 16134.

Xia, X., Ge, Y., Ge, F., Gu, P., Liu, Y., Li, P., & Xu, P. 2024. MAP4 acts as an oncogene and prognostic marker and affects radioresistance by mediating epithelial–mesenchymal transition in lung adenocarcinoma. *Journal of Cancer Research and Clinical Oncology*, 150(2), 88.

Yale Medicine. n.d. Non-Small Cell Lung Cancer: Symptoms, Diagnosis and Treatment. Yale Medicine.

Yang, J., Lin, M., Zhang, M., Wang, Z., Lin, H., Yu, Y., Zheng, Q., Chen, X., Wu, Y., Yao, Q., & Li, J. 2023. Advanced Glycation End Products' Receptor DNA Methylation Associated with Immune Infiltration and Prognosis of Lung Adenocarcinoma and Lung Squamous Cell Carcinoma. *Genetics Research*, 2023, 1–18.

Yang, Q., Zhang, H., Wei, T., Lin, A., Sun, Y., Luo, P., & Zhang, J. 2021. Single-Cell RNA Sequencing Reveals the Heterogeneity of Tumor-Associated Macrophage in Non-Small Cell Lung Cancer and Differences Between Sexes. *Frontiers in Immunology*, 12, 756722.

Zhang, Y., Wu, T., Li, F., Cheng, Y., Han, Q., Lu, X., Lu, S., & Xia, W. 2022. FGF19 Is Coamplified With CCND1 to Promote Proliferation in Lung Squamous Cell Carcinoma and Their Combined Inhibition Shows Improved Efficacy. *Frontiers in Oncology*, 12.