2020

Identifying mutational landscape and inflammatory pathways associated with lung squamous premalignant lesion progression

BOSTON UNIVERSITY

SCHOOL OF MEDICINE

Thesis

IDENTIFYING MUTATIONAL LANDSCAPE AND INFLAMMATORY PATHWAYS ASSOCIATED WITH LUNG SQUAMOUS PREMALIGNANT LESION PROGRESSION

by

JULIA CAMASSOLA BREDA

B.S., Lindenwood University, 2017

Submitted in partial fulfillment of the

requirements for the degree of

Master of Science

2020

Approved by

First Reader

Sarah Mazzilli, Ph.D.
Assistant Professor of Medicine

Second Reader

Jennifer Beane, Ph.D.
Assistant Professor of Medicine

Third Reader

Elizabeth Duffy, M.A.
Assistant Professor of Pathology and Laboratory Medicine

ACKNOWLEDGMENTS

I would like to express my deep gratitude to Dr. Sarah Mazzilli, my research supervisor, for her guidance, encouragement and valuable knowledge on this research work. I would also like to thank Dr. Jennifer Beane and Elizabeth Duffy, my thesis readers, for their advice, assistance and useful critique.

I would also like to extend my thanks to the Mazzilli Lab, Roxanna Pfefferkorn, Darren Chiu and Tachira Pichardo, for their help with laboratory protocols and instrument handling.

IDENTIFYING MUTATIONAL LANDSCAPE AND INFLAMMATORY PATHWAYS ASSOCIATED WITH LUNG SQUAMOUS PREMALIGNANT LESION PROGRESSION

JULIA CAMASSOLA BREDA

ABSTRACT

Lung cancer is the leading cause of death among all cancer types, and a subtype of lung cancer known as lung squamous cell carcinoma (LUSC) is the second most predominant subtype accounting for ~30% of all lung cancer cases. Lung cancer is typically diagnosed at advanced stages, resulting in poor prognosis with a 5-year survival rate of 19% between 2008 and 2014 (Siegel et al., 2019). The lung cancer genome has extensively been profiled in part by The Cancer Genome Atlas (TCGA), identifying mutational signatures and revealing the contributions of tumor microenvironment. TCGA work has enabled the development of biomarkers and therapeutics for both identifying and treating advanced lung cancer. However, due to late diagnosis, the prognosis of lung cancer still remains poor. Therefore, understanding the genomics and immune microenvironment of premalignant lesions (PMLs) that precede the development of LUSC is crucial in order to improve the ability to diagnose disease early and identify intervention strategies to enhance the prognosis of lung cancer. As part of this thesis project, my objective is to validate, in both cancer cell lines and a carcinogenic mouse model, previously identified genes and their correlations with important immuno-regulatory pathways, as well as their impact in the progression or regression of PMLs.

The Lung Pre-Cancer Genome Atlas (PCGA) has begun to achieve the aforementioned goals of understanding the early molecular events in the process of lung squamous carcinogenesis (Beane et al., 2019. In the prior study from our group, bronchial PMLs were obtained through bronchoscopy procedures from a high-risk screening cohort at Roswell Park and subject to a molecular and genomic profiling (Beane et al., 2019). Beane et al. identified four molecular subtypes of PMLs via transcriptional profiling, where one subtype, the proliferative subtype, of PMLs were found to be enriched for bronchial dysplastic lesions. Additionally, within samples in the proliferative subtype, it was demonstrated that immunoregulatory pathways, such as interferon signaling and antigen presentation, were suppressed in the progressive/persistent proliferative PMLs.

Further computational analyses of the lung PCGA transcriptional data has identified that Glutathione S-transferase P (GSTP1) is up-regulated in PMLs that remain stable as dysplastic lesions or progress into advanced dysplastic lesions, suggesting that there is an anti-correlation between the expression of GSTP1 and immunological responses. Understanding the relationship between GSTP1 and its downstream effectors is important, as targeting GSTP1 may be a strategy to stimulate an immunological response in PMLs to drive a regression rather than progression towards invasive carcinoma. In order to establish the relationship between GSTP1 and immune-regulatory pathways such as the STAT3 and NF-KB pathways, we identified baseline expression of the target genes in cell lines all genes were consistent with the computational findings. Subsequently, the knockdown of GSTP1 demonstrated a significant decrease in the expression of GSTP1 but no change in STAT3 and NF-KB. Based on these results, interleukins such as IL-8 and IL-

13 were used to stimulate immunological responses and expected to cause an increase in expression of STAT3 and NF-KB pathways. However, upon 48, 24 and 4 hours of IL-8 and IL-13 stimulation, STAT3 and NF-KB pathways did not present changes in expression.

Furthermore, in effort to establish a viable research model for studying premalignant lesions of LUSC, we utilized the NTCU mouse model, a carcinogenic model that has been previously identified to mimic squamous lung carcinogenesis due to smoking in humans. The histology of the PMLs triggered by the NTCU treatment in mice resembles the carcinogenesis observed in human lungs, in which lesions progress from squamous metaplasia to increasing grades of dysplasia. Previous studies have suggested that carcinogenic models share mutational burden and specific gene mutations with humans opposed to genetically engineered mice (Westcott et al., 2015). As part of my work, I aimed to establish the viability of the NTCU mouse model as a representative model of study for PMLs in humans by analyzing the histology of lung lesions in mice at 8, 12 and 16 weeks of NTCU treatment. Furthermore, I established methods for isolating DNA and RNA from the tail and lungs of both control and NTCU treated mice for whole exome sequencing (WES) and RNA sequencing in addition to use for validation of my GSTP1 finding in an *in vivo* setting. Our goal was to show similarities between genomic changes of PMLs from the NTCU mice and human PMLs in order to establish the NTCU model as a means to further study changes identified in the progression of human disease.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

CCLE	Cancer Cell Line Encyclopedia
CIS	Carcinoma in Situ
GST	Glutathione S-transferase
GSTP1	Glutathione S-transferase P 1
H&E	Hematoxylin & Eosin
IHC	Immunohistochemistry
IL-8	Interleukin 8
IL-13	Interleukin 13
LUAD	Lung Adenocarcinoma
LUSC	Lung Squamous Cell Carcinoma
NSCLC	Non-Small Cell Lung Cancer
NTCU	N-nitroso-trischloroethylurea
NF-KB	Nuclear Factor Kappa B
PAH	Polycyclic Aromatic Hydrocarbon
PCGA	Pre-Cancer Genome Atlas
PL	Piperlongumine
PML	Premalignant Lesion
SCLC	Small Cell Lung Cancer

TCGA	The Cancer Genome Atlas
TMA	Tissue Microarray
TSNA	Tobacco-Specific Nitrosamine
WES	Whole Exome Sequencing

INTRODUCTION

Lung Cancer

Cancer is the second leading cause of death in the United States, with one fourth of cancer deaths attributed to lung cancer (Siegel et al., 2019). According to the American Cancer Society, in 2019, around 142,670 patients died from lung cancer, which accounts for a higher number than those that die from breast, colon and prostate cancers combined. Lung cancer is divided into two main subtypes: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC). Where NSCLC accounts for 85% of all cases of lung cancer in the United States (Duma et al., 2019). NSCLC develops in epithelial compartments of the lung, where the most predominant subtypes are adenocarcinoma, squamous cell carcinoma and large cell carcinoma. Adenocarcinoma accounts for 40% of all lung cancer cases, squamous cell carcinoma accounts for 30% of all lung cancer cases and large cell carcinoma accounts for 5 to 10% of all cases (Duma et al., 2019).

The risk factors for NSCLC include age, exposure to carcinogens including cigarette smoking, second-hand smoking, exposure to asbestos and radon, and family history of lung cancer (de Groot et al., 2018). Current smokers are at a significantly higher risk of developing NSCLC compared to never smokers. The duration and cigarette smoked per day are used to calculate pack years, where higher pack years have increased risk (Janjigian et al., 2010). Former smokers also remain at higher risk for developing lung cancer compared to never smokers, regardless of pack years and how long ago smoking cessation has happened (Inoue□Choi et al., 2018).

Squamous Cell Carcinoma of the Lung

Lung squamous cell carcinoma (LUSC), also known as epidermoid carcinoma (*Squamous Cell Lung Cancer*, 2014), arises from squamous epithelial cells near the central bronchus (*Non-Small Cell Lung Cancer Treatment (PDQ®)–Health Professional Version*, 2019), and the location of the tumor can determine the symptoms, which can be presented as trouble breathing, coughing, chest pain, blood in the sputum (*Squamous Cell Lung Cancer*, 2014). LUSC is strongly linked to smoking, even more so than other NSCLC subtypes, because primary tumors of LUSC are usually located in central location of the bronchus, where in insult of smoking is the strongest, meanwhile subtypes such as adenocarcinoma and large cell carcinoma arise in the peripheral bronchi (Morabia & Wynder, 1991).

Carcinogens originating from smoking tobacco, such as polycyclic aromatic hydrocarbons (PAHs) and tobacco-specific nitrosamines (TSNAs) (Minna et al., 2002), have been previously observed to have the capacity of causing mutations, such as chromosomal deletions and single nucleotide variations in the epithelial layer of the respiratory tract (Walser et al., 2008). However, exposure to tobacco carcinogens alone is not a determinant factor for developing lung cancer, and therefore, genetic predisposition and effectiveness of one's immune system status also come into play (Singhavi et al., 2018). Tobacco carcinogens, more specifically TSNAs, can form DNA adducts, causing the double helix structure of the DNA to be disrupted (Singhavi et al., 2018). These DNA

adducts can occur at locations in the genome that encode for genes that are important for controlling cell cycle, such as tumor suppressor genes like TP53 (Singhavi et al., 2018).

Early diagnosis of LUSC is challenging, as the ability to detect precancerous lesions or early lung cancer is limited and the molecular changes are not well understood or characterized, limiting the diagnosis of the disease before symptoms have been observed. Although advances are being made in screening for some types of lung cancer, most lung cancer lesions remain detected at advanced stages. The National Lung Screening Trial (NLST) has observed a 20% reduction in lung cancer mortality by the use of low-dose CT scans, which shows that screenings for lung cancer prevention have been adopted (The National Lung Screening Trial Research Team, 2011). However, a significant number of these low-dose CT scans also yielded false positive results, causing the use of this method to fall short in its purpose. According to the NLST, approximately 57% of NSCLC cases are diagnosed at advanced stages, with only 8% of those being eligible for surgical removal (Chudgar et al., 2015). This is in part due to the fact that the identification of lung cancer at advanced stages where the cancer has metastasized renders treatments challenging.

LUSC becomes an invasive carcinoma upon spread to the nearest tissue, and possibly spreading to distinct organs. Invasive carcinomas of LUSC are characterized, and distinguished from other lung cancers, by the presence of keratin pearls, which are formed by the keratinization of the atypical cells (Suarez & Knollmann-Ritschel, 2017). Moreover, there are four variants of LUSC: clear cell, small cell, papillary and basaloid (Perez-Moreno et al., 2012). The clear cell variant occurs because of the abundant presence of glycogen within the cytoplasm (Edwards & Carlile, 1985), and this variant can be observed

at any histological subtype (Perez-Moreno et al., 2012). The small cell variant is characterized by the proliferation of basal bronchial stem cells with some areas of squamous differentiation (Perez-Moreno et al., 2012). Sometimes, the small cell variant of LUSC can mimic a small cell lung cancer (SCLC), seen that areas with squamous cell differentiation are very small. The papillary variant is confined to the bronchus wall, and its growth can cause an obstruction of the bronchus (Perez-Moreno et al., 2012) (Afridi et al., n.d.). Lastly, the basaloid variant is, histologically, very similar and hard to distinguish from the small cell variant, therefore, it also presents proliferation of basal bronchial stem cells and the formation of a basaloid pattern (Perez-Moreno et al., 2012).

TCGA of LUSC

Extensive research has been done and, therefore, allowed for the characterization of important pathways, genes and proteins that are altered in LUSC, contributing to the identification of feasible therapeutic targets for LUSC at the invasive carcinoma stage. The Cancer Genome Atlas (TCGA) has identified the genomic and epigenomic landscape of LUSC based on the analysis of DNA copy number, somatic mutations, mRNA sequencing, mRNA expression, promoter methylation and whole exome sequencing from tumor samples of patients with stage I through stage IV LUSC (Cancer Genome Atlas Research Network, 2012). The TCGA was able to determine that there are many similarities between LUSC and lung adenocarcinoma, however, mutations on EGFR and KRAS are rarely observed in LUSC. Furthermore, researchers observed that an amplification on

chromosome 3q was notably unique to LUSC, and that somatic copy number alterations are observed at a greater magnitude in LUSC (Cancer Genome Atlas Research Network, 2012). The most common mutation, observed in almost all tumors analyzed by the TCGA, was a somatic mutation in TP53 (Cancer Genome Atlas Research Network, 2012). Genes such as SOX2, CCND1, MYC, WHSC1L1/FGFR1 and EGFR were observed to be significantly amplified in LUSC (Campbell, Alexandrov, et al., 2016). In the whole exome sequencing analysis, the TCGA identified a somatic mutation rate of 8.1 mutations per megabase for LUSC, and which they determined to be considerably higher than other types of cancer such as breast cancer (1.0 per megabase) and colorectal carcinoma (3.2 per megabase) (Cancer Genome Atlas Research Network, 2012). Moreover, pathways responsible for cell cycle control, oxidative stress, squamous cell differentiation and apoptosis, such as CDKN2A/RB1, NFE2L2/KEAP1/CUL3, SOX2/TP63/NOTCH1 and PI3K/AKT, respectively. Amplification of the microRNA MIR205 has also been characterized as unique to LUSC when compared to other subtypes of NSCLC, and therefore, it has been used to identify and distinguish LUSC (Campbell, Alexandrov, et al., 2016). In addition, it was also noted that inactivating somatic mutations in the HLA-A gene, observed in LUSC tumors, results in immune suppression (Cancer Genome Atlas Research Network, 2012), and therefore, contributes to the capacity of cancer cells to escape from immune checkpoints and proliferate.

Premalignant Lesions of LUSC

The identification and understanding of the molecular characteristics that enable the progression of premalignant lesions (PML) of LUSC is crucial for enhancing detection and importantly establishing methods to intervene in the progression towards invasive carcinoma. These precancerous lesions are observed before the tumor becomes invasive. The histology of premalignant lesions is thought to be stepwise in nature where pseudostratified bronchial epithelial layer becomes increasingly less ordered. The histological progression includes hyperplastic and metaplastic lesions that can progress to low-grade dysplastic lesions, including mild to moderate dysplasia and then on to high-grade lesions, which is include severe dysplasia and carcinoma in situ (CIS) (Merrick et al., 2016). Basal cell hyperplasia is the first response of the bronchial epithelial to carcinogens, which can progress to squamous metaplasia. The metaplastic lesion resembles normal epithelial tissue (Figure 1A) in a way that basal layer is still distinguishable, however, the pseudostratified columnar morphology of the epithelium will begin to show characteristics of a squamous epithelium morphology (Rigden et al., 2016) (Figure 1B). Low-grade lesions, or mild and moderate dysplasia, are characterized by a mild to moderate change in the architecture of the epithelial layer and cellular atypia. Mild dysplasia mainly affects the apical one third of the epithelium layer, whereas moderate dysplasia is identified by increased cellular atypia in the apical two thirds of the epithelial layer. (Ishizumi et al., 2010) (Figure 1C & 1D). Moderate dysplasia is followed by severe dysplasia, in which the basal layer and apical layer are no longer distinguishable from one another (Figure 1E). In severe dysplasia, cellular atypia becomes more significant. When progressed through CIS,

the epithelial layer is fully changed, and the cells look fully differentiated with variations in their shape, nuclei, hyperchromatism and nucleus: cytoplasm ratio also becomes affected, causing a nuclear pleomorphism (Ishizumi et al., 2010)(Khieu & Butler, 2020) (Figure 1F).

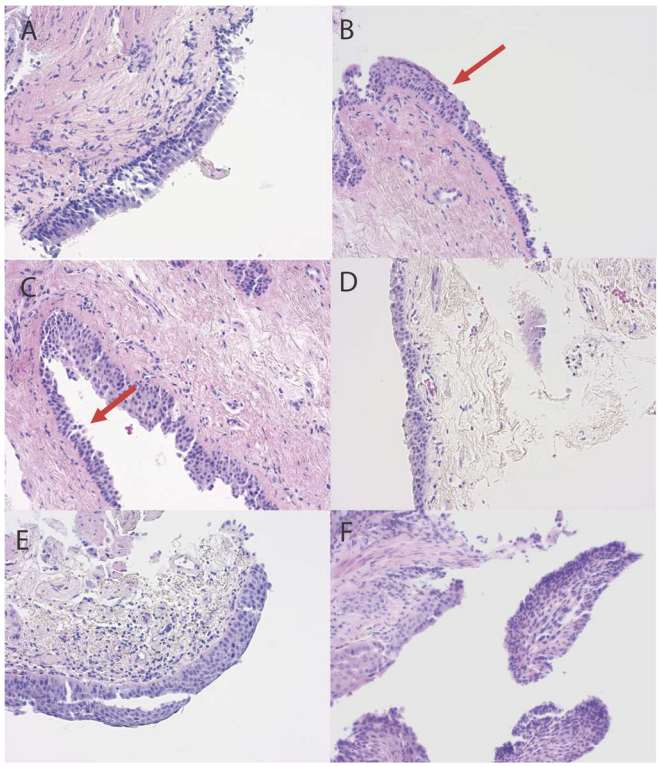

Figure 1: Histology of Premalignant Lesions in Human Lung. A: Histology of normal epithelial; **B:** Histology of squamous metaplasia stained with H&E; **C:** Histology of mild dysplasia; **D:** Histology of moderate dysplasia; **E:** Histology of severe dysplasia; **F:** Histology of CIS (Representative images from Pre-Cancer Genome Atlas)

The identification of CIS lesions in patients at risk for LUSC is becoming necessary in order to allow the diagnosis of LUSC at early stages. Previous studies have attempted to

identify and understand the trajectory and fate of CIS lesions, in order to make it possible, in the future, for the development of therapeutics and/or screening methods for early stages of LUSC. CIS lesions have been observed to go through many of the same hallmarks of cancer that LUSC does (Teixeira et al., 2019). However, many molecular differences between CIS lesions and invasive LUSC are also observed, and therefore, identifying them is of crucial importance in order to understand how the fate of these lesions can be determined early on.

Teixeira et al. performed studies that looked into establishing the genetic and epigenetic differences of CIS lesions that either regress to normal epithelium or progress to invasive carcinoma. They observed that the biopsies from a cohort of patients with PMLs, obtained at 4 and 12 weeks of screening, presented genomic differences, such as differential gene expression and DNA methylation sites, in CIS lesions that were either benign or progressed to invasive carcinoma (Teixeira et al., 2019). Regressive CIS lesions resemble normal bronchial epithelium, and appear to have less somatic mutations, less substitutions and less mutational signatures when compared to progressive CIS lesions (Teixeira et al., 2019). Furthermore, progressive lesions are highly associated with chromosomal instability, and many cancer-related pathways seem to be upregulated in these types of lesions. Understanding the molecular differences between these two CIS lesions can lead to the understanding of how tumorigenic progression of LUSC happens. In previous studies, researchers identified mutational signatures that were upregulated in progressive CIS lesions but were not upregulated in regressive CIS lesions, which were ACTL6A, ELAVL1, MAD2L1, NEK2 and OIP5 (Carter et al., 2006). Moreover, the

NKX2-1 gene was identified to be hypermethylated and downregulated in progressive lesions when compared to regressive lesions. Similarly, to what is observed in invasive lesions of LUSC, progressive CIS lesions also show overexpression of the SOX2 gene (Teixeira et al., 2019). Understanding what mutational signatures are exclusively observed in progressive CIS lesions gives a great insight into how, in the future, the progression of precancerous lesions of LUSC can be predicted. However, the previously mentioned epigenetic and genetic differences observed in these CIS lesions were only analyzed in the context of molecular changes of the epithelium, whereas immune microenvironment was not taken into account.

Pre-Cancer Genome Atlas

Driven by a need to understand premalignant lesions of cancer and, consequently, enable early identification of premalignant lesions (PMLs), the lung Pre-Cancer Genome Atlas (PCGA) had been established. This lung PCGA atlas initially consisted of the characterization of molecular changes in precancerous lesions, including profiling the genomic and transcriptomic changes and is now expanding to further characterizing the epigenetic, immune microenvironment and adding single cell characterization and spatial localization. Through the PCGA our groups are working to comprehensively shed light on the role of the immune microenvironment that contributes to the progression of PMLs to advanced stages of cancer (Campbell, et al., 2016). The lung PCGA has focused on PMLs of LUSC as they are accessible via bronchoscopy. Unlike LUSC, lung adenocarcinoma

(LUAD) is more commonly located in the distal airways, and therefore, reaching PMLs of LUAD via bronchoscopy is challenging due to possible damage of bronchioles and alveoli.

Recently published work by our group defined four transcriptional molecular subtypes of PMLs using endobronchial biopsies (Beane et al., 2019). The four subtypes of PMLs for LUSC: the normal-like subtype, secretory subtype, inflammatory subtype, and proliferative subtype differentiate from one another by differences in both immune processes and epithelial characteristics that were associated with gene modules or groups of genes with shared biological function. The normal-like subtype of PMLs showed an increase in the expression of genes involved in pathways associated with cilium/ciliagenesis, and there was significant decrease in the expression of genes that are involved in inflammation, lymphocyte and leukocyte regulation, antigen processing and presentation (Beane et al., 2019). The secretory subtype of PMLs was mostly associated with the upregulation of genes in cilium/ciliagenesis pathways, similar to what was observed in the normal-like subtypes, and downregulation of genes linked to protein translation. Immune cells and goblet cells seem to be upregulated in the secretory subtype, and this subtype is more commonly enriched in current smokers (Beane et al., 2019). The inflammatory subtype had increased expression of genes involved in RNA processing, and this is the most common subtype of PMLs observed in former smokers (Beane et al., 2019). The inflammatory subtype, a more abnormal subtype than normal-like and secretory, showed an upregulation in the expression of genes expressed by immune cells and it also had a downregulation of ciliated cells (Beane et al., 2019). Lastly, the proliferative subtype showed an increase in the expression of genes that are involved in energy metabolism as

well as cell cycle pathways. Marker genes for basal cells are overexpressed in the proliferative subtype, meanwhile, marker genes for club cells and ciliated cells are down-regulated (Beane et al., 2019).

As expected, the normal-like subtype showed a downregulation in CD45 cells (Beane et al., 2019). Furthermore, the four subtypes of PMLs were associated with LUSC subtypes defined in the TCGA, where the proliferative subtype was demonstrated to have high concordance with LUSC classical subtype, meanwhile, secretory and inflammatory subtypes of PMLs have a high concordance with LUSC secretory subtype (Beane et al., 2019). Moreover, the samples that were classified as the proliferative subtype that tended to regress from dysplasia to a lower grade lesion like normal or hyperplasia had higher expression of the immune Module 9 when compared proliferative PMLS that tend to progress/persist (Beane et al., 2019). Module 9 consists of a set of genes that are involved in immune responses, such as interferon signaling, antigen processing and presentation (Beane et al., 2019), and therefore, a higher expression of these genes will result in a more accurate performance and response from the innate and adaptive immunity, resulting in the progression of a PML into a normal-like epithelium.

On-going work in our group, is defining the mutational landscapes of LUSC PMLs, where mutations in genes such as NOTCH1, FAT1, TP53 and ARD1A are highly mutated in both PMLs (J. Campbell). Additionally, further analyses of the transcriptional data from the lung PCGA identified eight immune phenotypes in the PCGA PMLs, in which three differentiate regressive proliferative PMLs from progressive proliferative PMLs (C. Merestein, based on immune phenotypes in Thorsson et al). These immune phenotypes

associated with progression/regression are: interferon signaling, antigen presentation, and B- cells, and are somewhat overlapping with the previous gene modules identified by Beane et al. C. Merenstein has suggested that these immune models, in PMLs that progress from mild/moderate dysplasia towards high grade dysplasia, are downregulated, suggesting that these immunologic processes are suppressed (Figure 2) (Unpublished work).

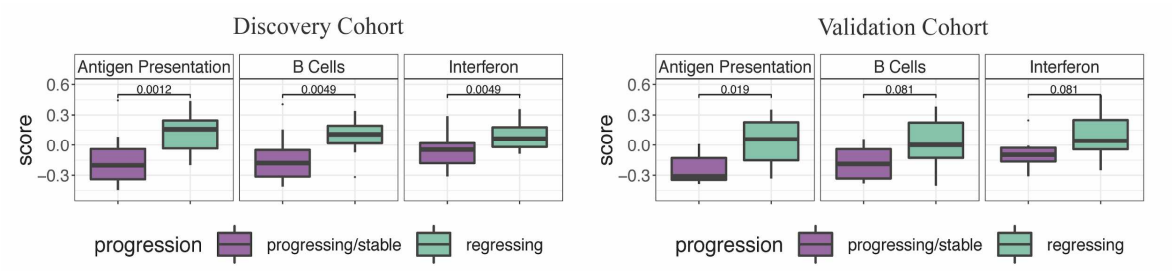

Figure 2: Immune phenotypes across progressive/persistent and regressive PMLs. Immune phenotypes across endobronchial biopsies from Roswell Comprehensive Cancer Center subjects, in which the Discovery Cohort (DC) samples (n = 190) were obtained between 2010 and 2012 and the Validation Cohort (VC) samples (n = 105) were obtained between 2012 and 2015 (Beane et al., 2019). Progressive/persistent PMLs (DC n = 44, VC n = 20) presented lowered scores for all three immune phenotypes, whereas regressive PMLs (DC n = 30, VC n = 18) presented higher scores. The boxes in the plot represent the progressive/persistent and regressive samples in purple and green, respectively, and how they are associated with the three immune phenotypes. The black line within the boxes represent the median score for the samples. In the Discovery Cohort plot, regressing lesions present higher medians for all three immune phenotypes when compared to

progression/persistent lesions, suggesting that these pathways are more likely upregulated in regressing lesions than in progressive/persistent lesions. Similar results were observed in the Validation Cohort plot, however, regressive lesion samples presented more dispersed scores throughout all the samples, which can be observed by taller green boxes in the plot, when compared to the VC plot. In general, for both DC and VC plots, progressive/persistent lesions presented a lower median for scoring of immune phenotypes, suggesting that the pathways are downregulated in these types of lesions. Figure generated by C. Merenstein, computational analyst (Beane/Lenburg/Spira Laboratories).

A mediation analysis correlating the association between gene expression, three immune phenotypes and proliferative lesion outcome (progressive/persistent versus regressive) identified genes that may be modulating the immune phenotypes. The mediation analysis identified a strong relationship between the gene GSTP1 and two of the three immune models identified in proliferative subtype PMLs. The relationship suggests that the upregulation of GSTP1 may contribute to the regulation of important immunological pathways such as the STAT3 and NF-KB pathways, resulting in a decreased immunologic response that contributes to the progression of these lesions (Figure 3).

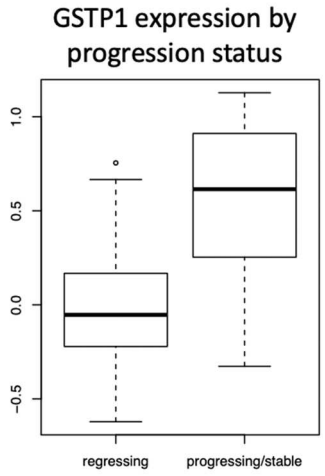

Figure 3: Expression of GSTP1 in regressive and progressive lesions. The regressing proliferative lesions show a lower expression of GSTP1 (on the left), which is reflected by a short box with a median line located at a lower scoring on the y-axis. Meanwhile, progressive/stable proliferative lesions have a higher expression of GSTP1 (on the right), with a tall box that reflects more dispersed scoring values throughout all samples, and with a median line much higher on the y-axis when compared to regressing lesions. The plot reflects the use of a linear mixed effects model controlled for TIN, batch and smoking (n = 295, p = 0.002). Figure generated by C. Merenstein, computational analyst (Beane/Lenburg/Spira Laboratories).

Immunoregulatory Pathways in Premalignant Lesions

The innate immune system is the first line of defense of a host against pathogens/ infected cells. Innate immune cells can recognize molecular patterns of pathogens/infected cells, known as PAMPs (pathogen-associated molecular patterns), through cell-surface

receptors known as cellular pattern recognition receptors (PPRs). The binding of a PAMP to a PPR promotes the production of cytokines and activation of B-cells (Lousberg et al., 2010). The adaptive immune system is, however, more complex and has the ability to develop an immunological memory that can serve as a backup immune protection for the host against malignant cells once the innate immunity has kicked in (Pandya et al., 2016). This immunologic memory allows the immune system to be aware of non-self antigens, therefore, allowing for immune cells to quickly respond against foreign antigens.

Immune subtypes or phenotypes, are specialized events in which the immune system can activate immune-specific cells and responses against non-self antigens. Interferon signaling is a first responder immune subtype, in which the production of interferons (IFNs) by the pathogen/infected cell can be recognized by the innate immunity, through the PPR of macrophages and dendritic cells, triggering the production of IFN by these cells (Ivashkiv & Donlin, 2014). The exposure of innate immunity cells to IFNs are suggested to trigger antigen presentation and production of cytokines and chemokines (Ivashkiv & Donlin, 2014). In the lung, interleukins such as IL-13 and IL- 8 are cytokines and chemokines, respectively, that are observed to be produced and released in response to inflammation. The expression of IL-13 is suggested to be increased in NSCLC, with the highest expression observed in LUSC in comparison to other subtypes of NSCLC (Pastuszak-Lewandoska et al., 2018). Chemokines, such as IL-8, were observed to be secreted from bronchial epithelial cells in response to cigarette smoking (Mortaz et al., 2011). Antigen processing/presentation is also an immune phenotype, in which antigens are first presented to MHC class I or II molecules by antigen-presenting cells, such as

dendritic cells or macrophages. This event allows MHC molecules to recruit and present antigens to naive T cells, either $CD4^+$ or $CD8^+$ cells. Upon antigen presentation, these naive T cells will differentiate and undergo clonal expansion (Cruse et al., 2004, Jain & Pasare, 2017).

Cancer cells have the ability to evade immune checkpoints, and therefore, proliferate without the innate and adaptive immunity recognizing them as non-self. Lowered interferon signaling and antigen processing/presentation suggests a decrease of immune responses in cancer cells. Immunoregulatory pathways, such as the STAT3 and NF-KB pathways, are crucial for the production of cytokines that promote appropriate immune responses against cancer cells. The blockade of these pathways can, therefore, allow for cancer cells to proliferate and continue to escape from immune responses, causing lesions to progress towards invasive carcinoma.

Glutathione S- transferase-P (GSTP1) is an enzyme that under normal circumstances functions in metabolizing xenobiotics, such as detoxifying electrophiles and free radicals that are toxic for both cells and genome (Xia et al., 1996). GSTP1 is an important enzyme from the GST superfamily located in the lung (Gu et al., 2014). GSTP1 is composed of two dimers, with an N-terminal and C-terminal domain found in each dimer (Allocati et al., 2018). The N-terminal domain is responsible for binding glutathione (GSH) in the G-site, which is located within the N-terminal (Allocati et al., 2018). The C-terminal domain contains the H-site, that along with the N-terminal, shapes the binding site for co-substrates of GSTP1 (Allocati et al., 2018). Furthermore, higher amounts of GSTP1 are

usually observed during the S and G_2 phases of the cell cycle (Xia et al., 1996), which correspond to phases in which the cells are being replicated and actively proliferating.

Analysis of GSTP1 expression in lung cell lines suggests it is anti-correlated with immunological processes. Cancer cell lines such as SW900 and H2170 along with immortalized cell lines such as the AALEs have shown an upregulation of the GSTP1, with a consequential downregulation of genes and pathways that GSTP1 interacts with, such as STAT3 and NF-KB (Figure 4) in sequencing data from the CCLE (Cancer Cell Line Encyclopedia) as well as data generated in our group.

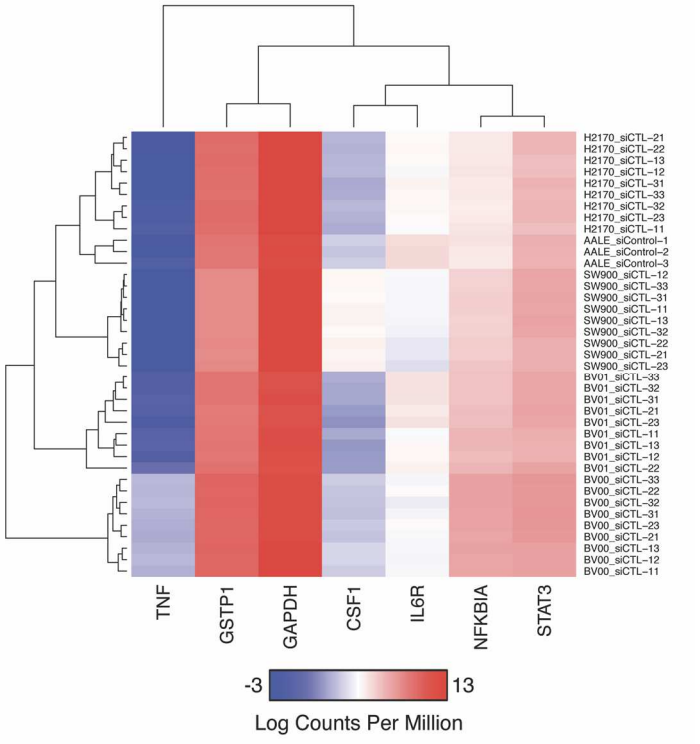

Figure 4: Expression of GSTP1 and target genes in cell lines. The heatmap depicts absolute values of gene expression of GSTP1, STAT3, NF-KB, IL6R, CSF1 and TNF in four different cell lines: SW900 and H2170, which are cancer cell lines; AALE, an

immortalized cell line; and BVO1, a human bronchial epithelial cell line profiled by RNA seq. Figure generated by C. Merenstein, computational analyst (Beane/Lenburg/Spira Laboratories).

The STAT3 pathway is activated once STAT3 is phosphorylated in response to specific cytokines and growth factors. The phosphorylated STAT3 can trigger the expression of many downstream genes, such as the colony stimulating factor 1 (CSF-1), that play roles in cell growth and apoptosis (Hillmer et al., 2016). GSTP1 is thought to interact with the STAT3 pathway by inhibiting the phosphorylation of STAT3, which consequently can no longer enter the nucleus and transcriptionally activate its downstream targets. As previously mentioned, CSF-1, or Colony-stimulating Factor 1, is one of the genes thought to be affected by GSTP1 and has been found to be one of the genes that is most anti-correlated to GSTP1. Furthermore, the NF-KB pathway is also affected by the GSTP1 expression. The NF-KB pathway regulates transcription, cell survival and cytokine production (Liu et al., 2017). Moreover, NF-KB is bound to its phosphorylated inhibitor, known as IkBa, which once phosphorylated will release NF-KB enabling its nuclear entrance to activate its targets, and IkBa will consequently get degraded by proteasomes in the cytoplasm (Karin, 1999). GSTP1 is thought to inhibit this pathway by interacting with TRAF2, or TNF-receptor-associated factor 2, which is an upstream factor of IkBa. Upon binding to TRAF2, GSTP1 inhibits the activation of IKK, which would subsequently phosphorylate IkBa (Jones et al., 2016). However, without the phosphorylation of IkBa, NF-KB can no longer detach from IkBa and enter the nucleus to transcriptionally activate

its downstream targets. One of the main targets of NF-KB is the tumor necrosis factor alpha (TNF-alpha), that is significantly downregulated once GSTP1 is perturbing the regulatory pathway, affecting its secretion by macrophages and monocytes as a red flag for an immunocompromised status (Idriss & Naismith, 2000).

NTCU - A Carcinogenic Mouse Model

Developing a pre-clinical model for precancerous lesions is of extreme importance in order for researchers to fully understand the process in which carcinogenesis happens in LUSC, and how every pathway behind every cellular and molecular function affects either the progression or regression of these lesions. Consequently, understanding and identifying the fault in every main biological process that contributes to tumor development will allow for the development of biomarkers and therapeutics that can detect/treat LUSC at such early stages.

Mouse models have been extensively used for the purpose of several diseases. Previous research has used mouse models for lung cancer studies given the similarities observed between human and the mouse lungs. The NTCU mouse model is a chemically induced model where treated mice develop premalignant lesions, such as squamous metaplasia and dysplastic lesions, before progressing to an invasive carcinoma (Yamano et al., 2016). Such ordered progression of premalignant lesions allows for characterization of the pre-invasive phase of the disease. The N-nitroso-trischloroethylurea (NTCU) compound is a carcinogen that induces the formation of premalignant lesions in mice that

can eventually progress into LUSC. NTCU is usually applied onto the skin of the mice and it is dose and duration-dependent. Previous studies have analyzed the histology of both PMLs, in humans and mice, and have found that there is a resemblance. The NTCU mouse model has been previously reported to be the only carcinogenic mouse model that can promote dysplastic lesions in the lungs of the mice that very closely resemble those observed in human smokers (Ghosh et al., 2015). The histology of progression of tumor carcinogenesis in mice lung, upon NTCU treatment, is similar to that observed in humans, in which the carcinogenic provokes the start of a development of flat atypia, in which some cells in the epithelium become enlarged with a hyperchromatic nuclei, followed by low grade dysplasia, high grade dysplasia, which then progresses and becomes squamous cancer and, therefore, no longer considered a PML (Ghosh et al., 2015, Mazzilli et al., 2015). Our group and others have demonstrated oncogenic pathways that are altered in humans PMLs and the airway brushing of individuals with LUSC were also modified in mice under the NTCU treatment (Xiong et al., 2017). Pathways such PI3K/AKT and Myc were found to be activated before the development of the PMLs were observed, which has also been previously determined to happen in humans with LUSC (Xiong et al., 2017).

As previously discussed, detection and treatment of early stages of lung cancer remains scarce and more work is needed to fulfill the unmet needs, however, a lot of work has already been done and will continue to discover transcriptional, genomic and microenvironmental changes. As part of this project, we hypothesized that GSTP1 is anticorrelated with the expression of immunoregulatory pathways such as STAT3 and NF-

KB in PMLs, and hope to show that interfering with GSTP1 expression can potentially promote the upregulation of these immune pathways and boost immune responses in PMLs, leading these lesions towards regression. We have identified the baseline expression of our target genes (GSTP1, STAT3, NF-KB, CSF1 and TNF-alpha) in cancer cell lines and tissue microarrays from LUSC samples to demonstrate the suggested inhibitory relationship of GSTP1 and STAT3 and NF-KB pathways. We hypothesize that upon GSTP1 knockdown, immune response stimulation with interleukins, oxidative stress stimulation and in the presence of cigarette smoke condensate that GSTP1 would be downregulated and STAT3 and NF-KB pathways would become upregulated as a response to immune stimulation in cancer cell lines. We have studied the expression of the target genes in tissue microarrays of LUSC to explore the validity of the suggested relationship between GSTP1 and STAT3 and NF-KB pathways in samples coming from actual tumors. Lastly, we aimed to further validate the NTCU carcinogenic model as a viable model for the study of lung squamous PMLs, and bridge the studies done in cancer cell lines and tissue microarrays by exploring the anticorrelation of GSTP1 with immune pathways in lung tissue samples coming from NTCU treated mice.

METHODS

Cell Culture

Grade IV Lung Squamous Cell Carcinoma cell line, known as SW900 (ATCC) and lung squamous cell carcinoma cell line H2170 (ATCC) were cultured in RPMI 1640 Medium (Gibco) containing 10% FBS (Gibco) and 1% Penicillin-Streptomycin (ThermoFisher Scientific). The cells were cultured at 37 °C and 5% CO_2 in T75 flasks and passaged once were 80-90% confluent. At every passage, cells were detached from the flask with Trypsin:EDTA Solution 0.25% (Gemini Bio-Products) and washed with DPBS 1x (Gibco). A vial containing approximately $1x10^6$ cells of the immortalized cell line AALE (ATCC) was thawed and washed with DPBS for the sole purpose of mRNA isolation.

Cell Transfection Assay

SW900 cells, approximately $3x10^5$ cells, were seeded in a 6-well plate with RPMI Medium for 24 hours to reach 70 to 80 % confluence. After 24 hours, cells were transfected with GSTP1 siRNA #AM51331 (ThermoFisher Scientific). The siRNA stock solution was prepared by diluting siRNA in water, to yield in a concentration of 10µM. The transfection solution was prepared by incubating 3µL of siRNA diluted in Opti-MEM Reduced Serum medium (Gibco) to 9µL of Lipofectamine RNAiMAX Transfection Reagent (ThermoFisher Scientific), which was also previously diluted in Opti-MEM medium. After 250µL of the siRNA-lipid complex was added to each well, the 6-well plate was incubated

at 37 °C for 48 hours. Cells were prepared for miRNA extraction by detaching them from wells with trypsin:EDTA, and centrifuging with DPBS 1x (Gibco) to remove impurities and trypsin from cell pellets. For control source, SW900 cells were treated with Stealth RNAi™ siRNA Negative Control #1293512 (ThermoFisher Scientific). The negative control siRNA, 1.5µL of 20µM stock solution of siRNA, was incubated with 5µL of Lipofectamine RNAiMAX in Opti-MEM medium for 15 minutes. An amount of 500µL of negative control siRNA and Lipofectamine solution was added to each well. The 6-well plate was incubated at 37 °C for 48 hours.

Immune Response Stimulation Assay

The SW900 cells were treated with 10 ng/µL of IL-13 (Stemcell Technology) and 10 ng/µL of IL-8 (Sigma Aldrich), separately. The interleukins were diluted into 2 mL of RPMI Medium, and added to each well individually. The 6-well transfected plates containing either IL-13 or IL-8 were incubated for 4, 24 and 48 hours. The 6-well plates transfected with negative control siRNA were also treated with IL-13 and IL-8 at the same concentration and time as the GSTP1 siRNA transfected cells.

Oxidative Stress Stimulation Assay

The SW900 cells were seeded in a 6-well plate, with $3x10^5$ cells per well and incubated until they were 80-90% confluent. The cells were then treated with 50µM of 30% Hydrogen Peroxide (H_2O_2) (Fisher Scientific). The H_2O_2 was diluted in 2mL of RPMI Medium for each well of the 6-well plate, and incubation of the cells with H_2O_2

was carried out for 4 hours. The control variable was SW900 cells that received 2mL of RPMI Medium, per well.

Cigarette Smoke Condensate Assay

The SW900 cells were seeded in a 6-well plate for 24 hours to allow for 80-90% confluence. The cells were treated with 100µg/mL of cigarette smoke condensate (CSC) (Murty Pharmaceuticals). A volume of 2.5 µL of CSC was used for every mL of RPMI Medium, and a total of 2mL of diluted CSC in RPMI Medium was added to each well of the plate. The cells were incubated with CSC for 4 hours. The control cells were treated with 2.5µL of 96% Dimethyl Sulfoxide (DMSO) (Fisher Scientific) for every mL of RPMI Medium, and incubated for 4 hours.

RNA Isolation from Cell Pellets & Frozen Lung Tissue

After cells were ready for RNA isolation, they were treated with 700µL of QIAzol lysis reagent (QIAGEN) and 140 µL of Chloroform (Sigma Aldrich). The tubes were centrifuged for 15 minutes at 4°C and 13.2 rpm. The upper most layer of the trilayer was collected, around 350 µL, and treated with 100% ethanol. The cells were placed in an RNA spin column and went through a series of washes with buffers from the miRNease Mini Kit (QIAGEN). The RNA was eluted with 30 µL of UltraPure^{TM} Distilled water (Invitrogen). Once isolation was completed, 1.5 µL of the RNA samples were quantified in the Spectrophotometer Nanodrop.

RT-PCR Assay

The cDNA synthesis was performed with the use of the Transcriptor First Strand cDNA Synthesis Kit (Roche). The Master Mix prepared for the cDNA synthesis contained: RNA template to a final concentration of 1µg, 1µL of Anchored Oligo $(dT)_{18}$ Primer, 2µL of Random Hexamer Primer, 4µL of Transcriptor RT Reaction Buffer 5x, 2µL of dNTP mix, 0.5µL of Protector RNase Inhibitor and 0.5µL of Transcriptor Reverse Transcriptase, and PCR-grade water to add up to 20µL total per reaction tube. The cDNA synthesis reaction is done by the Veriti Thermal Cycler (Applied Biosystems). Tubes are incubated for 10 minutes at 25 °C, followed by 30 minutes at 55 °C, and the Transcriptase enzyme is then inactivated by heating the tubes to 85 °C for 5 minutes. The tubes are then chilled down to 4 °C.

The PCR reagents used along with cDNA were TaqMan Primers for GAPDH, GSTP1, STAT3, NF-KB, TNF alpha and CSF1 (ThermoFisher Scientific) and TaqMan Universal PCR Master Mix (ThermoFisher Scientific). The primers, Universal PCR Master Mix and cDNA template were loaded onto a 96-well PCR plate (Applied Biosystems), and reactions were run in triplicates, with each reaction volume was 20 µL. Each reaction contained 4µL of cDNA, 16µL of PCR-grade water and 4µL of TaqMan primer. The plate was run in the StepOne Real-Time PCR System (Applied Biosystems) with a standard ramp speed. The thermal cycling conditions were: One hold stage at 50°C for 2 minutes, followed by a second hold stage at 95°C for 10 minutes, followed by 40 cycles of 1 minute each at a temperature of 60°C.

The relative expression of genes ($2^{-\Delta\Delta CT}$), for the baseline expression studies, was calculated by using the ΔC_T method by employing GAPDH as the housekeeping gene. The $\Delta\Delta C_T$ of GSTP1 was calculated by subtracting the ΔC_T value for GSTP1 in the samples treated with GSTP1 siRNA from the ΔC_T value for GSTP1 in the samples treated with negative control siRNA (ΔC_T of treated samples - ΔC_T untreated sample). All the $\Delta\Delta C_T$ values for the remaining genes were calculated with the same formula. The relative fold expression, or fold-change, for the GSTP1 knockdown and interleukin stimulation assays, was calculated using the $2^{-\Delta\Delta CT}$ method, also normalizing for GAPDH.

Slides Preparation and Staining

Lungs were harvested from mice at 8, 12 and 16 weeks of NTCU treatment and control, and frozen for storage and further use. For slide preparation, frozen lungs were placed in Tissue Plus O.C.T Compound (Fisher HealthCare). Lung sections of 8 μm were cut with cryostat and placed in a Colorfrost/Plus Microscope Slide (Fisherbrand). The lung sections were stained with Modified Harris Hematoxylin (Richard-Allen Scientific) for 6 minutes, washed with Acid EtOH for up to 3 seconds, and subsequently counterstained with Eosin Y Stain, 0.25% (w/v) in 57% (v/v) alcohol (RICCA) for approximately 15 seconds. The slides were then washed in a series of ethanol dilutions (70%, 95% and 100%) for 3 minutes each, and further dehydrated in three different washes of Xylene for approximately 3 minutes. Once dehydration was completed, the slides were mounted with VectaMount Permanent Mounting Medium (Vector), and covered with Premium SUPERSLIP^{TM} coverslips (Fisher Scientific) for further microscopical analysis.

Immunohistochemistry protocol was performed on TMAs (Tissue Microarrays) from lung squamous cell carcinoma tumor samples, in which were prepared from 4 μm sections of freshly cut FFPE embedded tumor and stained on a Ventana Benchmark Ultra (Ventana Medical Systems). Onboard heat induced epitope retrieval and Optiview detection (Ventana Medical Systems) was used for visualization. The TMAs were stained with polyclonal and unconjugated antibodies sourced from rabbit. The primary antibodies used were: anti-GSTP1 antibody #HPA019869 (Sigma Aldrich); Phospho-STAT3 (Tyr305) antibody #9131 (Cell Signaling Technology); NF-KB p65 antibody #GTX107678 (GeneTex); anti-IKBKB antibody #HPA001249 (Sigma Aldrich); anti-M-CSF antibody #Ab9693 (Abcam) and TNF-alpha antibody #GTX110520 (GeneTex).

The staining intensity was scored on a grade ranging from 0 to 3, in which zero was determined to have no staining of the target antibody within the tumor tissue, and score 3 was considered to present the highest intensity of antibody staining.

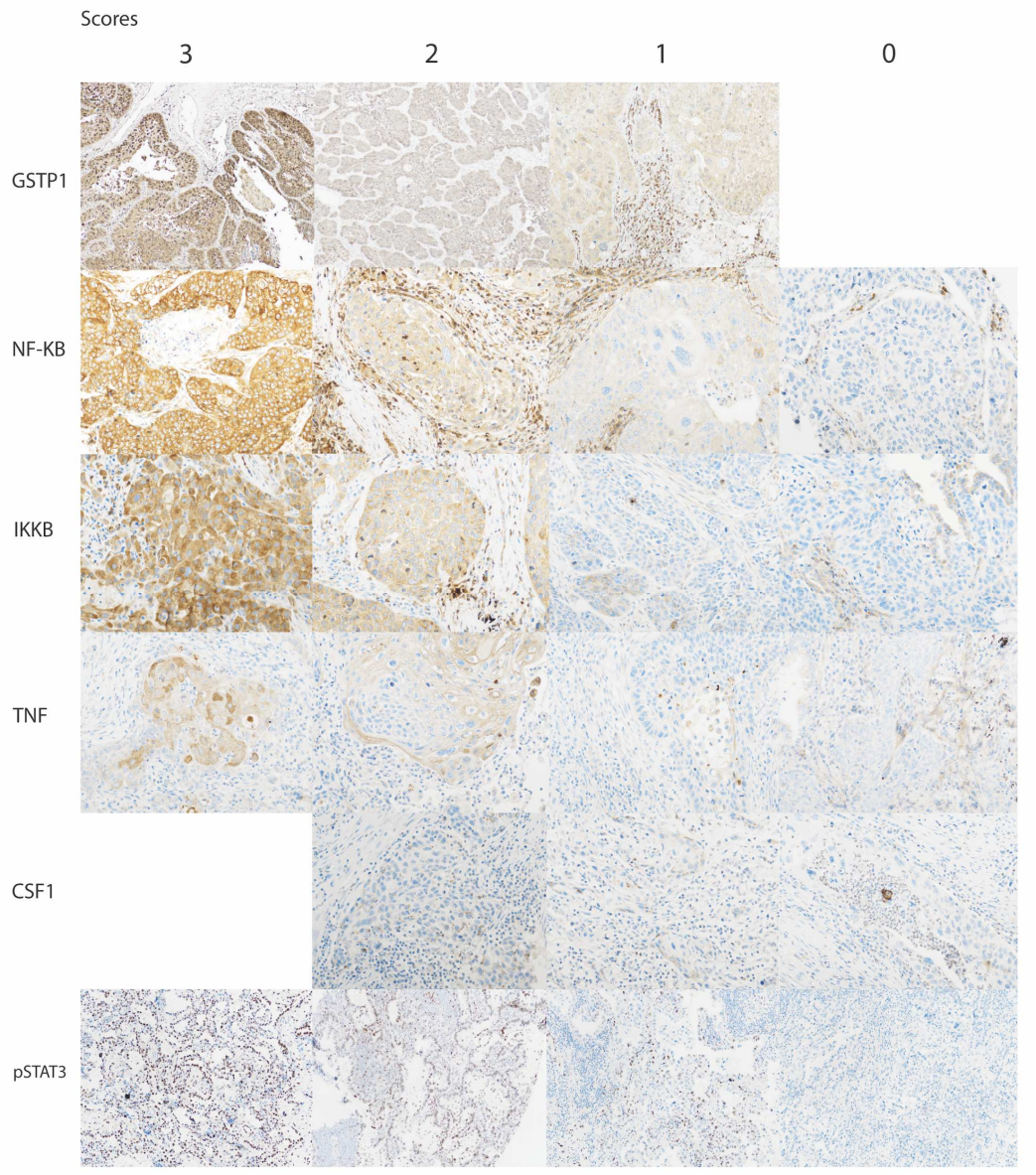

Figure 5: Staining Scores of Target Genes in Immunohistochemistry. Each staining was scored on a scale from 0 to 3, with 3 as the highest staining intensity observed in the analysis of the TMA slides. Some genes, such as GSTP1 and CSF1, did not present specific scoring grades throughout the TMA samples that were analyzed.

DNA and RNA Isolation from Frozen Tissue

Genomic material from both control and NTCU treated mice were extracted from frozen harvested lungs with the use of QIAMP DNA/RNA Kit (QIAGEN). Frozen lungs embedded in O.C.T Compound were sectioned at different thickness (20 and 25 μm) with the cryostat to yield in 80 and 100 μm curls. The tissue curls were isolated and purified from the remaining components of the lung tissue according to QIAMP DNA/RNA Kit Protocol. The lung curls were lysed with 600μL of RTL buffer (QIAGEN) and placed in an Allprepr DNA spin column (QIAGEN) and centrifuged for 30 seconds at full speed. The flow-through was used for RNA extraction. The DNA was further lysed with 20μL of proteinase K and 60μL of buffer AW1 (QIAGEN). The genetic material was further purified with a series of washes with Buffer AW1 and AW2 (QIAGEN). The elution was done with 60 and 30μL of UltraPure™ Distilled water. The RNA was lysed with 80μL of proteinase K, and vortexed with two washes of 100% ethanol. The RNeasy spin column (QIAGEN) was used, and 80μL of DNase I and buffer RDD stock solution was added to the RNA, incubated for 15 minutes. The RNA was further washed with buffer RPE (QIAGEN) and ethanol. The elution was done with 30μL of UltraPure™ Distilled water.

The tails from each mouse were used as an internal control, and genomic material was extracted by transforming the frozen tail into powdery consistency with the help of a mortar and pestle (Fisherbrand). The tails were then further lyzed and purified according to QIAMP DNA Mini Kit (QIAGEN) Protocol. The tails were individually placed into 180μL of buffer ATL (QIAGEN) and 20μL of Proteinase K (QIAGEN), and incubated at ~56°C for 10 minutes. Another 200μL of buffer AL was added and the samples were

incubated at ~70°C for 10 minutes. The samples were then vortexed with 200µL of 100% ethanol, and the mixture was placed in a QIAp mini spin column (QIAGEN). The DNA in the column was further washed with 500µL of AW1 and AW2 buffers (QIAGEN), separately. The DNA samples were eluted with 60µL of Ultra-Pure water (Gibco).

DNA and RNA Quantification and Qualification

The DNA samples coming from frozen tails and lungs, as well as RNA samples from frozen lungs were quantified with the use of a Spectrophotometer Nanodrop 2000C (Thermo Scientific). RNA samples were further analyzed for quality and integrity with the use of a BioAnalyzer (Agilent).

RESULTS

GSTP1 and Target Genes: Determining Baseline Expression on Cancer Cells and Immortalized Cell Lines

Previous computational analysis has demonstrated that GSTP1 is upregulated in certain cancer cell lines, as well as in immortalized bronchial epithelial cell lines. We, therefore, have established baseline gene expression of GSTP1, STAT3, NF-KB, CSF1 and TNF-alpha in the cell lines we are using: SW900 (LUSC), H2170 (LUSC) as well as AALE (bronchial epithelial) cells. We have used GAPDH, a housekeeping gene, as an endogenous control knowingly that its expression is fairly constant and high throughout

many different cell lines. The SW900 cell line, as expected, presented low C_T value for GSTP1, 21.43, and higher C_T values for STAT3, NF-KB, CSF1 and TNF-alpha (Table 1). The H2170 cell line presented a C_T value for GSTP1 similar to the one observed for SW900, 20.36. As for the other genes of interest, C_T values also were higher, confirming that as GSTP1 is upregulated, STAT3 and NF-KB pathways are downregulated (Table 1). Interestingly, the AALE cell line also had a C_T value for GSTP1, 19.92, that corresponded to what was expected from previous computational analysis (Table 1).

Altogether, GSTP1 and the other genes of interest presented an absolute expression at baseline close to what was established in computational analysis. More specifically, as observed in the gene expression heatmap (Figure 4), TNF-alpha expression is significantly lower for all cell lines, and TaqMan RT-PCR results have shown extremely high C_T values, suggesting that there is limited amplification for TNF-alpha (Figure 6). Similarly, CSF1 was observed to have extremely low expression in the H2170 cell line (Figure 6), and as observed through RT-PCR results, the C_T value for CSF1 in H2170 cells was 32.89, meanwhile in SW900 and AALE cells it was 27.72 and 28.93, respectively (Table 1). The $\Delta\Delta$CT values were not calculated for this part of the experiment, seen that there were no treatments/interventions done, and the focus was to determine baseline gene expression of the genes.

	H2170 Cells	SW900 Cells	AALE Cells
GAPDH	21.34 ± 2.00	20.59 ± 1.29	20.75 ± 0.20
GSTP1	20.36 ± 2.89	21.43 ± 1.39	19.92 ± 0.04
STAT3	25.19 ± 0.48	24.79 ± 0.59	24.98 ± 0.01
NF-KB	25.04 ± 0.45	25.48 ± 0.93	23.91 ± 0.06
CSF1	32.05 ± 0.87	27.29 ± 0.56	28.93 ± 0.03
TNF-alpha	33.89 ± 0.65	36.83 ± 0.13	32.59 ± 0.30

Table 1: C_T Values for Genes of Interest in H2170, SW900 and AALE Cells at Baseline. The table presents mean ± standard deviation of C_T values for each gene of interest (GAPDH, GSTP1, STAT3, NF-KB, CSF1 and TNF-alpha) in each different cell line of interest (H2170, SW900 and AALEs).

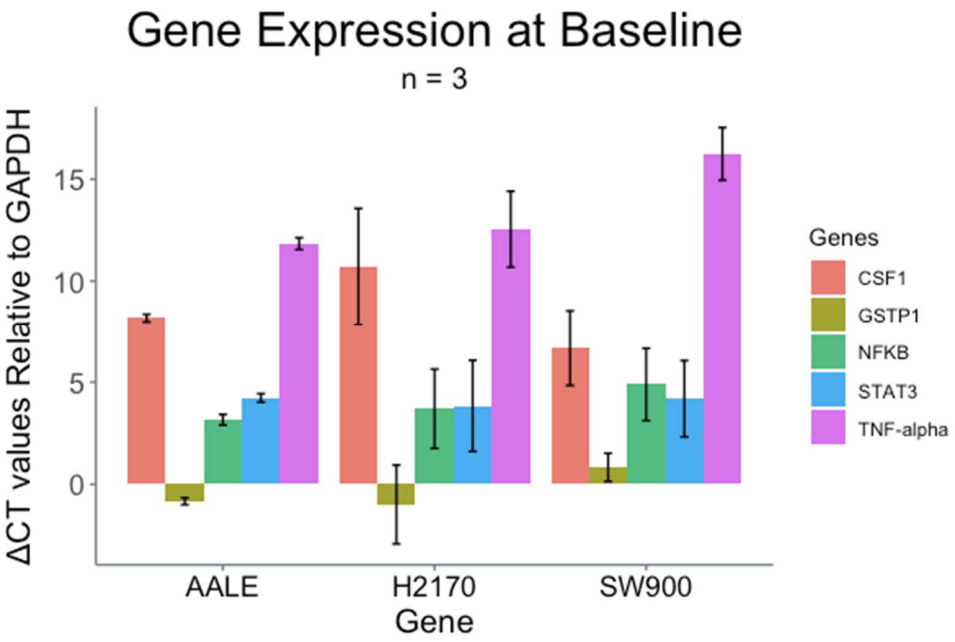

Figure 6: Gene expression of Target Genes at Baseline. The expression of the target genes was calculated by normalizing C_T values of genes to C_T value of GAPDH, yielding in ΔC_T values. High ΔC_T values means that the gene is downregulated in comparison to GAPDH. The opposite is also true, in which low or negative ΔC_T values means that the

gene is either expressed similarly or upregulated in comparison to GAPDH. The gene expressions by cell line (AALE, H2170 and SW900), and each gene of interest is assigned a different color. It is observable that each gene is expressed similarly in each one of the cell lines, with exception of CSF1 that seems to be more downregulated, with higher ΔC_T values, in H2170 cells.

GSTP1 Knockdown with siRNA and Stimulation of Immune Responses with IL-13 and IL-8

In effort to establish the role of GSTP1 in modulating the transcription of immune modulatory genes (STAT3, NF-KB, CSF1 and TNF-alpha) we have employed siRNA technology to knock down GSTP1. The transfection protocol for the knockdown of GSTP1 was done by incubating SW900 cells with GSTP1 siRNA for 48 hours. The knockdown yielded a C_T value of 28.73 for GSTP1 (Table 2). The efficiency of knockdown of GSTP1 was quite efficient with a 24-fold decrease in the expression of GSTP1 in the siRNA transfection assay when compared to SW900 cells treated with negative control siRNA (Table 2) (Figure 7B). However, the expression of STAT3, NF-KB, CSF1 and TNF-alpha remained similar in the negative control SW900 cells when compared to baseline expression values previously obtained (Table 1). In fact, there was a small decrease in fold-change and expression of STAT3, NF-KB and TNF-alpha in the GSTP1 siRNA transfected cells (Table 2) (Figure 7A & 7B).

	GAPDH	GSTP1	STAT3	NFKB	CSF1	TNF-alpha
Negative Control siRNA (C_T value)	21.11 ± 0.32	21.54 ± 0.41	24.28 ± 0.06	23.45 ± 0.08	25.27 ± 0.08	33.49 ± 0.15
GSTP1 siRNA Transfection (C_T value)	23.87 ± 0.11	28.73 ± 0.19	26.92 ± 0.13	26.69 ± 0.09	27.83 ± 0.11	36.61 ± 0.61
$\Delta\Delta C_T$		4.65	-0.12	0.48	-1.78	0.19
$2^{-\Delta\Delta CT}$ (Relative Expression)		0.04 ± 0.007	0.82 ± 0.023	0.73 ± 0.178	0.59 ± 0.502	0.73 ± 0.135
-1 / $2^{-\Delta\Delta CT}$ (Fold-Change)		-24.81	-1.21	-1.37	-1.67	-1.37

Table 2: C_T, $\Delta\Delta C_T$ and fold-change values for control and GSTP1 siRNA transfected SW900 cells. Cells transfected with siRNA for 48 hours, control sample was treated with negative control siRNA, and the treated sample was transfected with GSTP1 siRNA. $2^{-\Delta\Delta CT}$ values, or relative expression, are presented as mean ± sd.

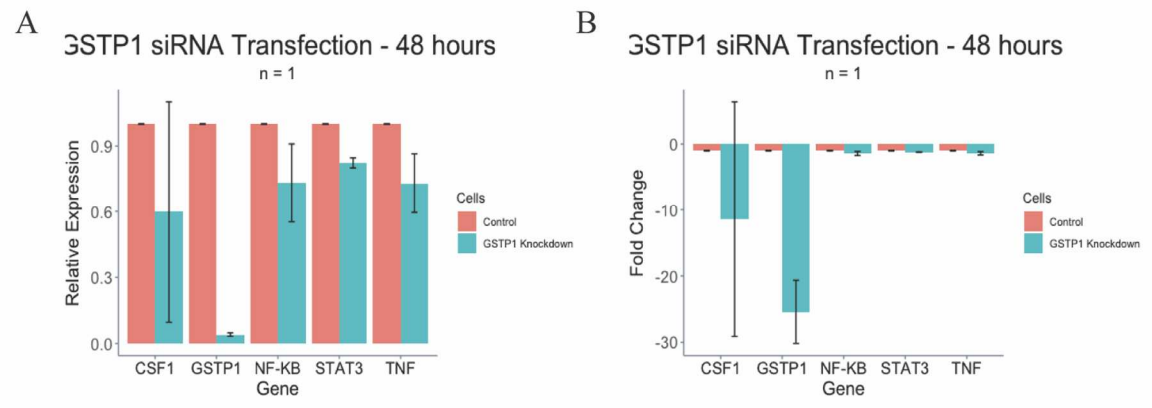

Figure 7: 48 hours GSTP1 siRNA Transfection in SW900 Cells. A: $2^{-\Delta\Delta CT}$ values for GSTP1, CSF1, NFKB, STAT3 and TNF-alpha of SW900 cells treated GSTP1 siRNA for 48 hours. Values are normalized for GAPDH. GSTP1 has a very low relative expression in comparison to GSTP1 expression in control cells, suggesting its successful knockdown. CSF1, NF-KB, STAT3 and TNF-alpha relative expressions are relatively lower in the absence of GSTP1. **B:** The fold-change values for each gene of interest after transfection

with GSTP1 siRNA for 48 hours. The fold-change values are normalized from GAPDH. GSTP1 and TNF-alpha had the greatest decrease in fold-change of all genes. CSF1, NF-KB and STAT3 showed a slight decrease in expression.

Once we observed that GSTP1 siRNA transfection successfully knocked-down GSTP1, but that genes such as CSF1 and TNF-alpha remained at low expression, we furthered the studies and stimulated the cells with inflammatory cytokines suggested to be increased in the airway with the addition of 10 ng/mL of IL-13 and IL-8 to the transfected cells (Pastuszak-Lewandoska et al., 2018, Mortaz et al., 2011). The interleukins were added to two separate 6-well plates. We have observed that after 48 hours of siRNA transfection + 10ng/mL of IL-13, STAT3, NF-KB, CSF1 and TNF-alpha expression remained relatively low when compared to control cells that did not have GSTP1 knockdown (Figure 8). We then observed the effects of IL-13 on the transfected cells for 24 hours, and we discovered that the treatment with IL-13 for 24 hours led to very similar results as the 48-hour treatment. There was a small difference in the decrease fold-change for all target genes, however, STAT3 and NF-KB pathways were still not activated. TNF-alpha was undetectable at qPCR for the IL-13 treatment for 24 hours (Figure 9).

As the responses from stimulation of immune responses with IL-13 for 24 and 48 hours in the SW900 cells did not result in upregulation of CSF1 and TNF-alpha, we then aimed to treat the cells with IL-13 for a shorter amount of time. After the transfection of the cells for 48 hours with either negative control siRNA or GSTP1 siRNA, we stimulated the cells with IL-13 for 4 hours. GSTP1 knockdown and treatment with IL-13 for 4 hours drove TNF-alpha to a significant down-regulation of its expression (Figure 10). The NF-

KB gene showed a small decrease in fold-change (Figure 10), however, this gene was expected to maintain its expression. Genes such as STAT3 and CSF1 genes maintained a fairly constant expression upon IL-13 stimulation in both negative control and GSTP1 knockdown (Figure 10).

	GAPDH	**GSTP1**	**STAT3**	**NFKB**	**CSF1**	**TNF-alpha**
Neg. Control + IL13 (C_T value)	24.20 ± 0.22	22.55 ± 0.58	24.80 ± 0.12	23.89 ± 0.14	26.65 ± 0.09	35.47 ± 0.29
GSTP1 siRNA +IL13 Transfection (C_T value)	20.88 ± 0.09	26.92 ± 0.13	24.73 ± 0.07	24.95 ± 0.06	26.16 ± 0.06	35.03 ± 1.84
ΔC_T (Neg. Control)		-1.66	0.60	-0.32	2.46	11.26
ΔC_T (GSTP1 siRNA + IL13)		6.04	3.85	4.07	5.28	14.15
$\Delta\Delta C_T$		7.69	3.25	4.39	2.82	2.88
$2^{-\Delta\Delta CT}$ (Relative Expression)		0.005 ± 0.001	0.12 ± 0.03	0.05 ± 0.01	0.14 ± 0.03	0.24 ± 0.241
-1 / $2^{-\Delta\Delta CT}$ (Fold-Change)		-148.19 ± 137.27	-9.69 ± 2.24	-21.24 ± 4.61	-7.19 ± 1.52	-10.19 ± 10.99

Table 3: Expression of Target Genes Upon IL-13 Stimulation for 48 Hours. Amplification and expression of target genes upon treatment with IL-13 for 48 hours. The negative control was transfected with siRNA with no homology to any genes and treated with IL-13 at same dosage and time as GSTP1 siRNA transfected cells. The negative control cells presented a slight downregulation in GAPDH, and GSTP1 was upregulated, unlike the GSTP1 siRNA transfected cells. Amplification values for remaining genes were fairly similar in both control and GSTP1 knockdown cells.

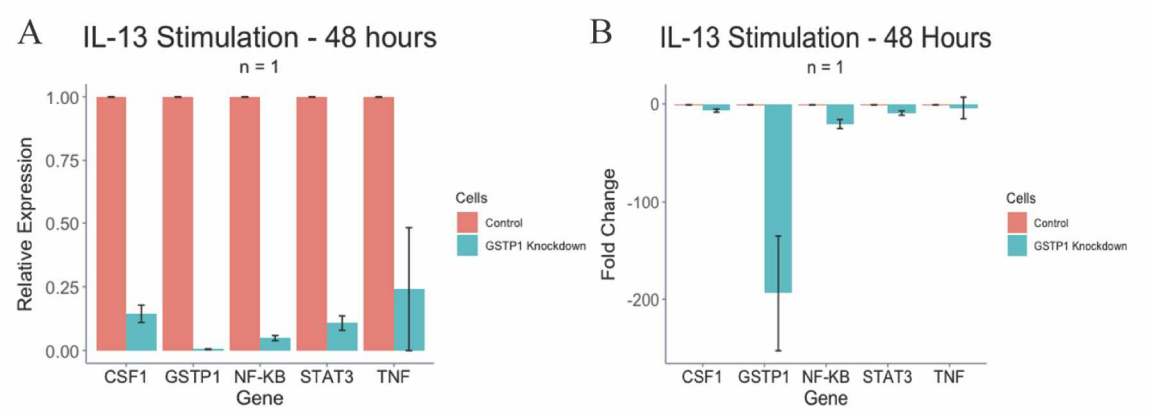

Figure 8: GSTP1 siRNA Transfection and IL-13 Stimulation for 48 hours. A: The mean ± standard deviation of $2^{-\Delta\Delta CT}$ values for all genes of interest, normalized for GAPDH. All genes presented relatively low expression values when compared to control cells. **B:** The fold change for each gene of interest after GSTP1 has been knocked down and IL-13 has been administered to stimulate an inflammatory response. STAT3, NF-KB, CSF1 and TNF-alpha presented a decrease in fold-change, when compared to negative control cells, under stimulatory presence of IL-13.

	GAPDH	**GSTP1**	**STAT3**	**NFKB**	**CSF1**	**TNF-alpha**
Neg. Control + IL13 (C_T value)	22.20 ± 0.12	22.13 ± 0.16	25.66 ± 0.09	24.24 ± 0.06	26.95 ± 0.02	36.40 ± 0.14
GSTP1 siRNA +IL13 Transfection (C_T value)	25.82 ± 0.14	29.66 ± 0.08	28.97 ± 0.03	28.04 ± 0.08	29.87 ± 0.08	-*
ΔC_T (Neg. Control)		-0.07	3.46	2.04	4.74	-*
ΔC_T (GSTP1 siRNA + IL13)		3.85	3.15	2.22	4.06	-*
$\Delta\Delta C_T$		3.92	-0.31	0.19	-0.69	-*
$2^{-\Delta\Delta CT}$ (Relative Expression)		0.07±0.02	1.25±0.22	0.64 0.31	1.6 ±0.15	-*
-1 / $2^{-\Delta\Delta CT}$ (Fold-Change)		-15.41 ± 3.87	-0.81±0.13	-1.97±1.29	-0.62±0.06	-*

Table 4: Stimulation of Immune Responses following treatment with IL-13 for 24 hours. The amplification and expression values resulting from RT-PCR. *TNF-alpha was not detected in the qPCR run, and therefore, does not have any amplification values for GSTP1 knockdown + IL13 cells. Expression values for TNF-alpha could not be calculated, seen that there was no amplification detected for any of the three replicates in the plate.

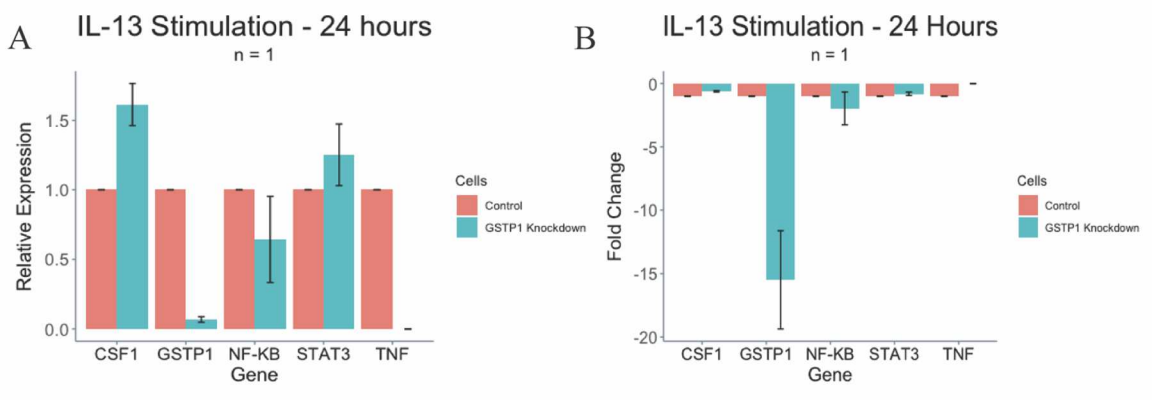

Figure 9: Stimulation with IL-13 for 24 hours. A: The mean ± sd of $2^{-\Delta\Delta CT}$ values of relative expression of target genes, once were normalized for GAPDH. Low relative expression value for GSTP1 is attributed to its knockdown with siRNA, the remaining genes remained with fairly constant expression in between control and GSTP1 knockdown. CSF1 and STAT3 present higher relative expression than control due to higher C_T value that was attributed to GAPDH in GSTP1 knockdown cells. TNF-alpha does not present any relative expression value in GSTP1 knockdown as the gene was not amplified in the qPCR run. **B:** The fold-change of target genes. The GSTP1 knockdown led to a decrease in fold change of GSTP1 to be higher than the other genes. However, genes such as CSF, which was expected to be upregulated, also present a small decrease in fold-change in expression.

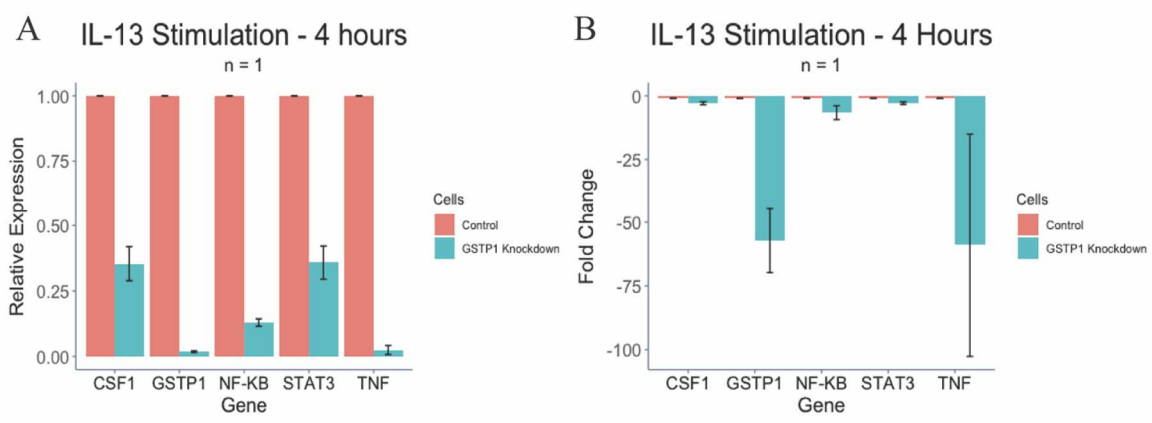

Figure 10: Stimulation with IL-13 for 4 hours. A: The graph represents mean ± sd of $2^{-\Delta\Delta CT}$ values of relative expression of target genes, once values were normalized for GAPDH. GSTP1 and TNF-alpha were the most downregulated genes, showing a lower relative expression when normalized to GAPDH. CSF1, NF-KB and STAT3 were not as amplified as GAPDH, however, their expression levels were not considerably low; **B:** Mean±sd fold-change of target genes. GSTP1 and TNF-alpha were the genes that were most down-regulated when analyzing negative control and GSTP1 knockdown cells.

We aimed to stimulate an immune response from the knockdown model of GSTP1 in SW900 by treating them 10ng/mL of IL-8. The 48 hour stimulation with IL-8 lead to very similar results as the IL-13 stimulation. We observed that GSTP1 was knocked down, with a decrease in the fold-change of 36.70 (Figure 11B). The STAT3 expression was relatively low in GSTP1 knockdown cells (Figure 11A). Genes such as TNF-alpha and CSF1, also had a lower relative expression than control cells, however, their decrease in fold-change was higher than STAT3 (Figure 11B). Furthermore, TNF-alpha and CSF1

showed a slightly higher decrease in fold-change (Figure 11B) when compared to treatment with IL-13 for 48 hours (Figure 8).

The SW900 cells were also stimulated with IL-8 for a 4 hour period, after the cells had already been transfected with either negative control siRNA or GSTP1 siRNA. We observed that, when cross-analyzing negative control versus GSTP1 knockdown, GSTP1 and TNF-alpha were significantly downregulated upon GSTP1 knockdown and IL-8 stimulation, whereas CSF1 and STAT3 maintained a fairly constant expression across both control and knockdown cells (Figure 12A).

	GAPDH	**GSTP1**	**STAT3**	**NFKB**	**CSF1**	**TNF-alpha**
Neg. Control + IL8 (C_T value)	23.30 ± 0.11	22.83 ± 0.4	25.92 ± 0.06	24.18 ± 0.3	26.81 ± 0.09	35.29 ± 0.59
GSTP1 siRNA +IL8 Transfection (C_T value)	22.56 ± 0.13	27.24 ± 0.22	25.29 ± 0.06	25.52 ± 0.04	27.53 ± 0.09	36.98 ± 0*
ΔC_T (Neg. Control)		-0.47	2.61	0.88	3.50	11.99
ΔC_T (GSTP1 siRNA + IL8)		4.68	2.74	2.96	4.97	14.42
$\Delta\Delta C_T$		5.16	0.12	2.08	1.46	2.43
$2^{-\Delta\Delta CT}$ (Relative Expression)		0.03±0.01	0.65±0.36	0.24±0.06	0.38±0.03	0.29±0*
$-1 / 2^{-\Delta\Delta CT}$ (Fold-Change)		-37.60±11.19	-2.10±1.58	-4.32±0.98	-2.68±0.19	-3.83±0*

Table 5: Expression of target genes following stimulation with IL-8 for 48 hours. Relative Expression and fold-change values are given as mean±sd values. Analyzing C_T values, all genes, with exception of GSTP1, had a small decrease in expression. *TNF-alpha does not have a standard deviation value, seen that only one of the triplicates in the PCR plate gave off an amplification C_T value.

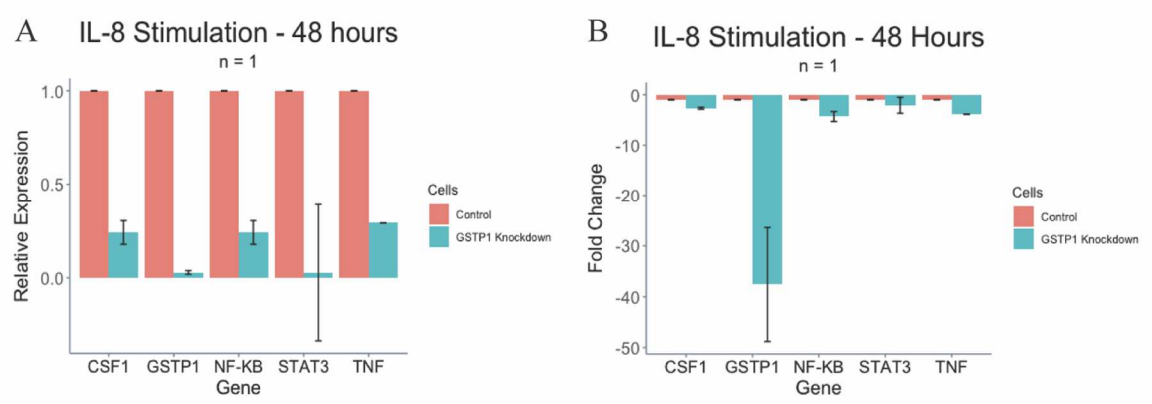

Figure 11: Stimulation with IL-8 for 48 hours. A: The mean ± sd of $2^{-\Delta\Delta CT}$ values of relative expression of target genes, once values were normalized for GAPDH. GSTP1 and STAT3 showed the lowest expression relative to control cells; **B:** The mean ± sd fold-change of target genes. GSTP1 was significantly knocked down, and NF-KB was slightly down-regulated as well.

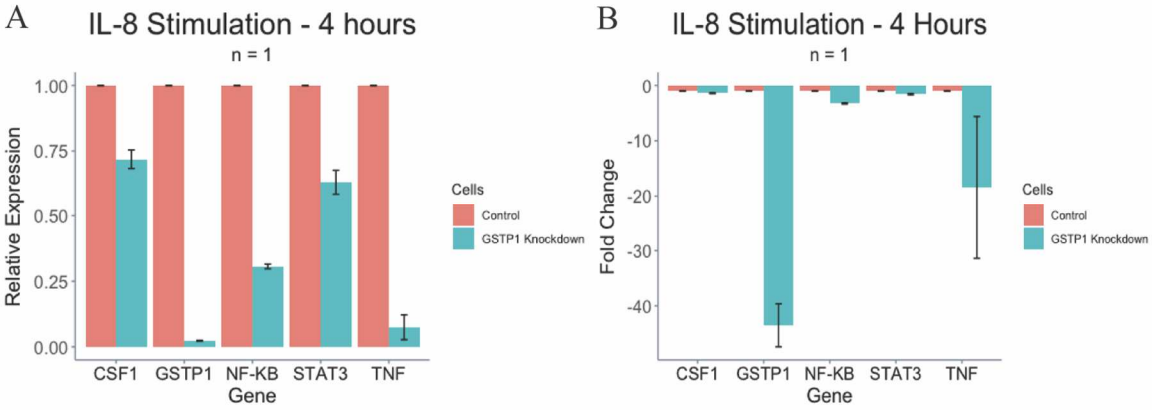

Figure 12: Stimulation with IL-8 for 4 hours. A: The mean ± sd of $2^{-\Delta\Delta CT}$ values of relative expression of target genes, once values were normalized for GAPDH; GSTP1 and TNF-alpha were significantly down-regulated in comparison to control cells; **B:** The mean±sd fold-change of target genes. GSTP1 and TNF-alpha are the genes with most

decrease in fold-change, whereas CSF1 and STAT3 were slightly downregulated, however, the amplification values were maintained constant. NF-KB was slightly downregulated after the GSTP1 knockdown.

Target Genes Response to Oxidative Stress and Cigarette Smoke Condensate

Observing that STAT3 and NF-KB pathways after stimulation of immune responses with interleukins upon GSTP1 knockdown were suggested to not be activated, thus to further investigate GSTP1's role as a possible inhibitor of these pathways. In this further analysis, we have induced oxidative stress in the cells using 50μM of hydrogen peroxide for 4 hours. The SW900 cells treated with H_2O_2 were not transfected with GSTP1 siRNA, therefore, GSTP1 expression on untreated cells was expected to present a high expression. After the 4 hours oxidative stress induction with H_2O_2 we observed that most genes maintained their relative expression levels in both treated and untreated cells, however, TNF alpha was considerably up-regulated on cells treated with H_2O_2 (Figure 13A). Furthermore, even though relative expression between untreated and treated cells was very similar, the fold-change of NF-KB and STAT3 were suggested to be significantly downregulated in treated cells (Figure 13B). TNF-alpha, in accordance with its increase in relative expression, showed an increase in fold-change when compared to untreated cells (Figure 13B).

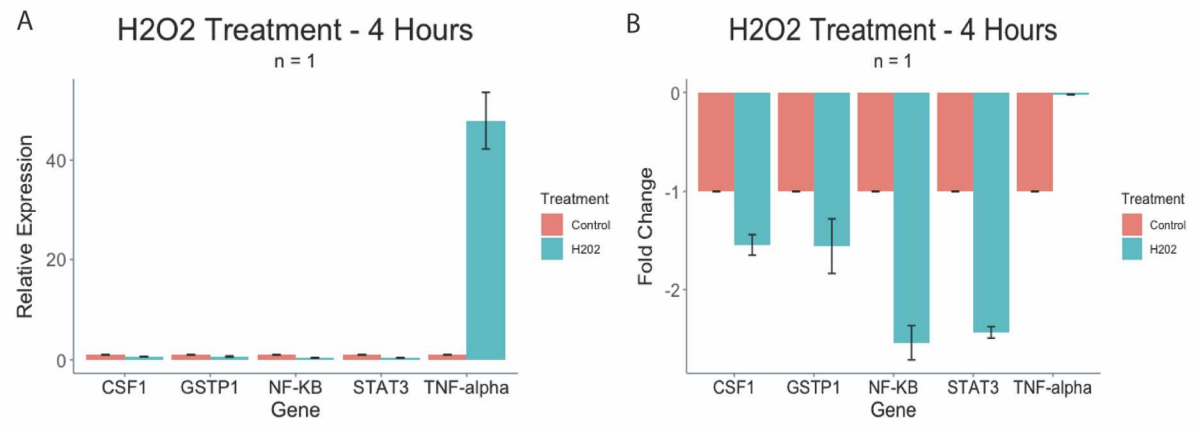

Figure 13: Induction of Oxidative Response with H_2O_2 in SW900 cells. A: mean ± sd relative expression of target genes. Control cells and H_2O_2 treated cells showed similar expression of most target genes, with exception of TNF-alpha, in which upon oxidative stress was upregulated. **B:** mean ± sd fold-change of target genes. The genes were downregulated, however, TNF-alpha showed the least decrease in fold-change.

	GAPDH	GSTP1	STAT3	NFKB	CSF1	TNF-alpha
Control Cells (C_T value)	22.59	22.82	25.89	25.65	29.32	36.00
H_2O_2 Treated Cells (C_T value)	21.44	22.29	26.03	25.84	28.79	29.28
ΔC_T (Control Cells)		0.22	3.29	3.05	6.72	13.40
ΔC_T (H_2O_2 Treated Cells)		0.85	4.58	4.39	7.35	7.83
$\Delta\Delta C_T$		0.63	1.28	1.34	0.63	-5.58
$2^{-\Delta\Delta CT}$ (Relative Expression)		0.66±0.12	0.41±0.01	0.39±0.03	0.65±0.04	47.89±5.65
$-1 / 2^{-\Delta\Delta CT}$ (Fold-Change)		-1.56±0.27	-2.43±0.06	-2.53±0.17	-1.55±0.10	-0.02±0.002

Table 6: Expression of Target Genes Following Induction of Oxidative Stress with H_2O_2. Relative Expression and fold-change values are given as mean±sd values.

As a method for analyzing GSTP1's possible anti-correlation with STAT3 and NF-KB pathways in smokers, we have treated SW900 cells with 100µg/mL of cigarette smoke condensate (CSC) for 4 hours. We observed that the effects of CSC on SW900 did not seem to affect genes such as GSTP1, NF-KB and STAT3. The genes presented similar relative expression in untreated and treated cells (Figure 14A) as well as similar fold-change values (Figure 14B). The CSF1 gene in treated cells was downregulated (Figure 14A), presenting the higher decrease in fold-change among all target genes (Figure 14B). TNF-alpha was suggested to be up-regulated in treated cells, presenting a higher relative expression than untreated cells, however, its fold-change was also decreased. Such effect can be attributed to the fact that one the three replicates on the 96-well PCR plate presented a considerably higher fold-change than the other two replicates, and that effect can be observed in the error bar plotted in the fold-change graph (Figure 13B).

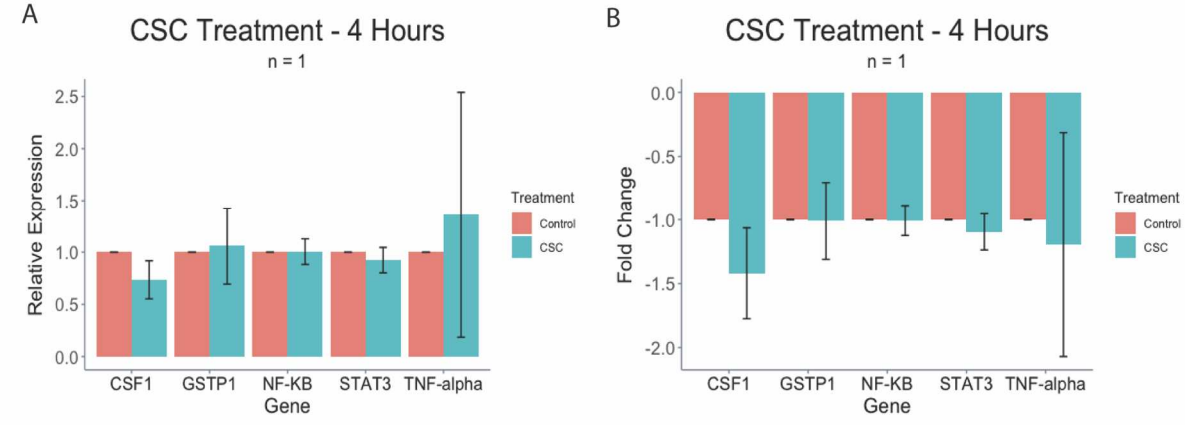

Figure 14: Effects of Cigarette Smoke Condensate on SW900 Cells. A: mean ± sd relative expression of target genes. **B:** mean ± sd fold-change of target genes upon addition of CSC for 4 hours.

	GAPDH	GSTP1	STAT3	NFKB	CSF1	TNF-alpha
Control Cells (C_T value)	21.87	22.51	25.56	25.61	29.42	35.07
CSC Treated Cells (C_T value)	22.02	22.62	25.84	25.76	30.05	35.15
ΔC_T (Control Cells)		0.63	3.69	3.74	7.55	13.19
ΔC_T (CSC Treated Cells)		0.59	3.81	3.74	8.02	13.12
$\Delta\Delta C_T$		-0.04	0.12	0	0.48	-0.07
$2^{-\Delta\Delta CT}$ (Relative Expression)		1.06±0.37	0.92±0.12	1.01±0.12	0.73±0.18	1.36±0.18
$-1 / 2^{-\Delta\Delta CT}$ (Fold-Change)		-1.01±0.30	-1.09±0.14	-1.00±0.12	-1.42±0.36	-1.19±0.88

Table 7: Expression of Target Genes Following Cigarette Smoke Condensate Treatment for 4 hours. Relative Expression and fold-change values are given as mean±sd values.

Staining of GSTP1 in Human Lung TMA

Tissue microarrays (TMAs) were used in effort to make a connection between gene expression and protein expression as we haven't yet done this in the cancer cell lines, in which we have conducted the majority of the experiments presented in this thesis. We aim to use this data to connect the experimental work with human lung tissue coming from squamous PMLs. In effort to analyze the expression, we have used immunohistochemistry (IHC) with antibodies for our genes of interest: GSTP1, pSTAT3, NF-KB, IKKB, CSF1

and TNF-alpha. In this analysis, we have also added IKKB, where the positive staining in the tumor tissue reflects that IKKB has not been degraded, and therefore, suggesting that NF-KB has not been translocated to the nucleus and activated. However, positive staining of NF-KB suggests that it has been activated and, therefore, translocated to the nucleus for the activation of its downstream targets. A positive staining of NF-KB should reflect a negative staining of IKKB, as both cannot be activated at the same time. Furthermore, positive staining of pSTAT3 reflects that the STAT3 has been phosphorylated and the pathway has been activated as STAT3 gets translocated into the nucleus, however, the absence of staining intensity of pSTAT3 reflects that no phosphorylation has taken place and STAT3 has not been translocated. Lastly, positive stainings for TNF-alpha and CSF1 are expected to be observed as positive stainings for NF-KB and STAT3, respectively, are present, suggesting that these genes have transcriptionally activated their downstream targets (TNF-alpha and CSF1).

The TMAs analyzed included stage I and II LUSC tumor cores from 101 patients. Each tissue section was stained for each of the antibodies of the target genes. We are utilizing a qualitative measure of staining intensities from 0 to 3, in which 3 corresponds to intense staining for the specific antibody within the tumor. When analyzing overall expression of each target gene, we have observed that GSTP1 predominately scored a grade of 2, accounting for 54% of all TMAs samples, with only 1% of the samples scoring a grade 0 (Figure 15A). The overall expression of pSTAT3 showed that 44% of samples were a score 0, and 44% were a score 1, however, none of the samples had a score of 3 (Figure 15B). The expression of NF-KB was interestingly similar to scores of GSTP1, and

44% of tumor tissue samples with a score of 2, 26% with a score of 3, 25% with a score of 1, and 5% with a score of 0 (Figure 15C). The overall expression of IKKB, as the gene being in direct correlation with NF-KB, presented inverse results as NF-KB, with the majority percentage of the tissue samples being either a score 0 or 1, with 34% and 50%, respectively (Figure 15D). Furthermore, overall expression of downstream target of NF-KB pathway, TNF-alpha, had low scores across all tumor samples, in which 30% of samples were scored 0 and 54% received a score of 1 (Figure 15E). Lastly, the expression of CSF1 was not expressed across most of the TMAs, leading to 94% of the samples receiving a score of zero (Figure 15F).

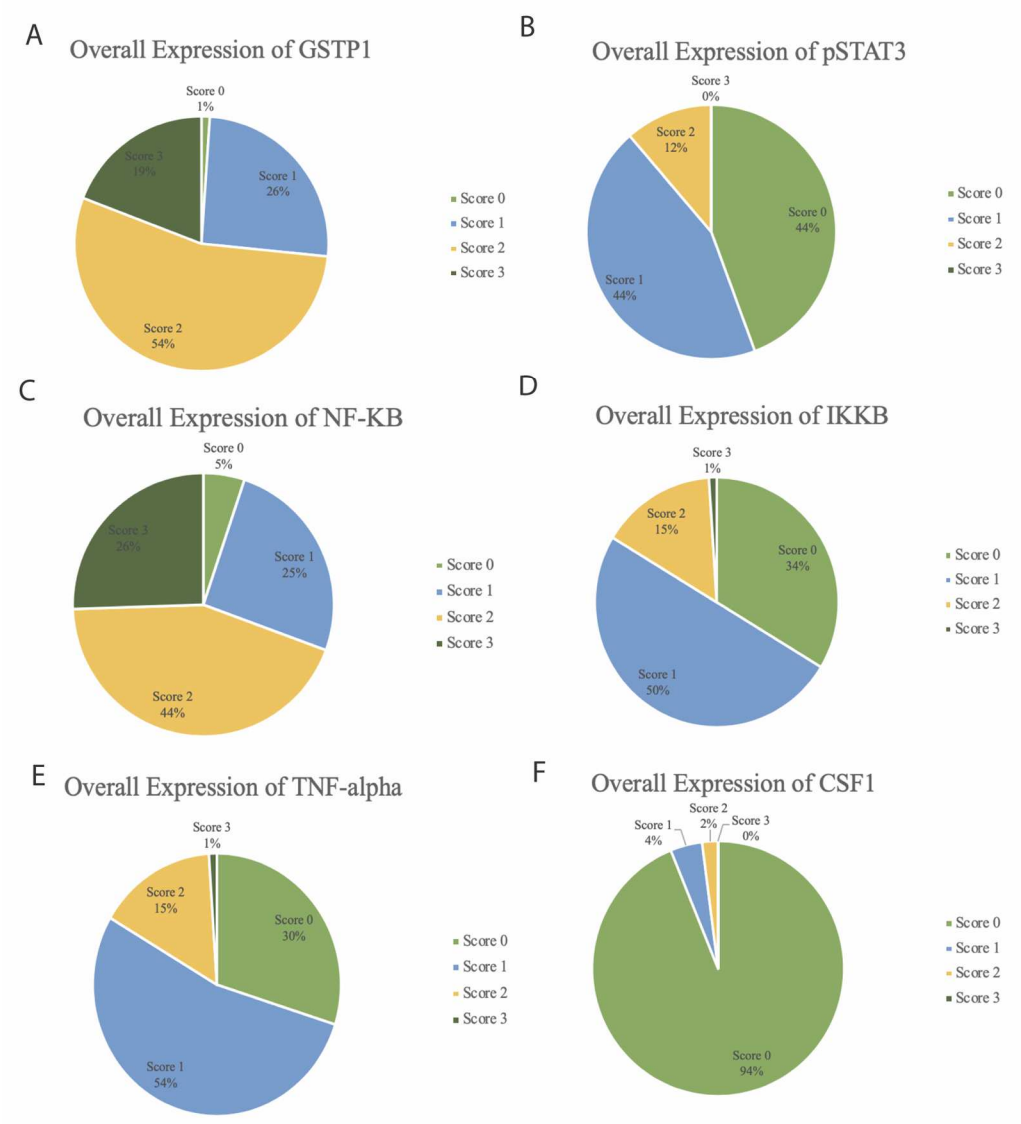

Figure 15: Overall Expression of Target Genes in Squamous Cell Carcinoma Tumor Tissue. A: Overall expression of GSTP1, depicted as the percentage of each scoring grade across 101 SCC tumor samples. **B:** Overall expression of pSTAT3. **C:** Overall expression of NF-KB. **D:** Overall expression of IKKB. **E:** Overall expression of TNF-alpha. **F:** Overall expression of CSF1.

In an attempt to assess the expression of each marker in the context of GSTP1 expression, we examined how each marker was expressed for each qualitative GSTP1 score. We have observed that tumor tissue sections that had high scores for GSTP1 (scores of 2 and 3), had low corresponding expression of pSTAT3, suggesting that higher expression of GSTP1 in the tissue may be associated with low pSTAT3 (Figure 16A). Whereas, high expression of GSTP1 appeared to be associated with increasing expression of cytoplasmic NF-KB (Figure 16B). This suggests that the expression of GSTP1 may not suppress expression of cytoplasmic NF-KB, however the absence of nuclear localization suggests inactivity. The staining intensity of IKKB in most tumors had low expression compared to NF-KB (Figure 16F), suggesting that IKKB released NF-KB for its nuclear translocation and that IKKB is being degraded (Figure 16C). Moreover, the comparison between GSTP1, TNF-alpha and CSF1 remained similar to what has been observed in our previous experiments, in which higher expression of GSTP1 is correlated with lower expression of TNF-alpha and CSF1 (Figure 16D & 16E).

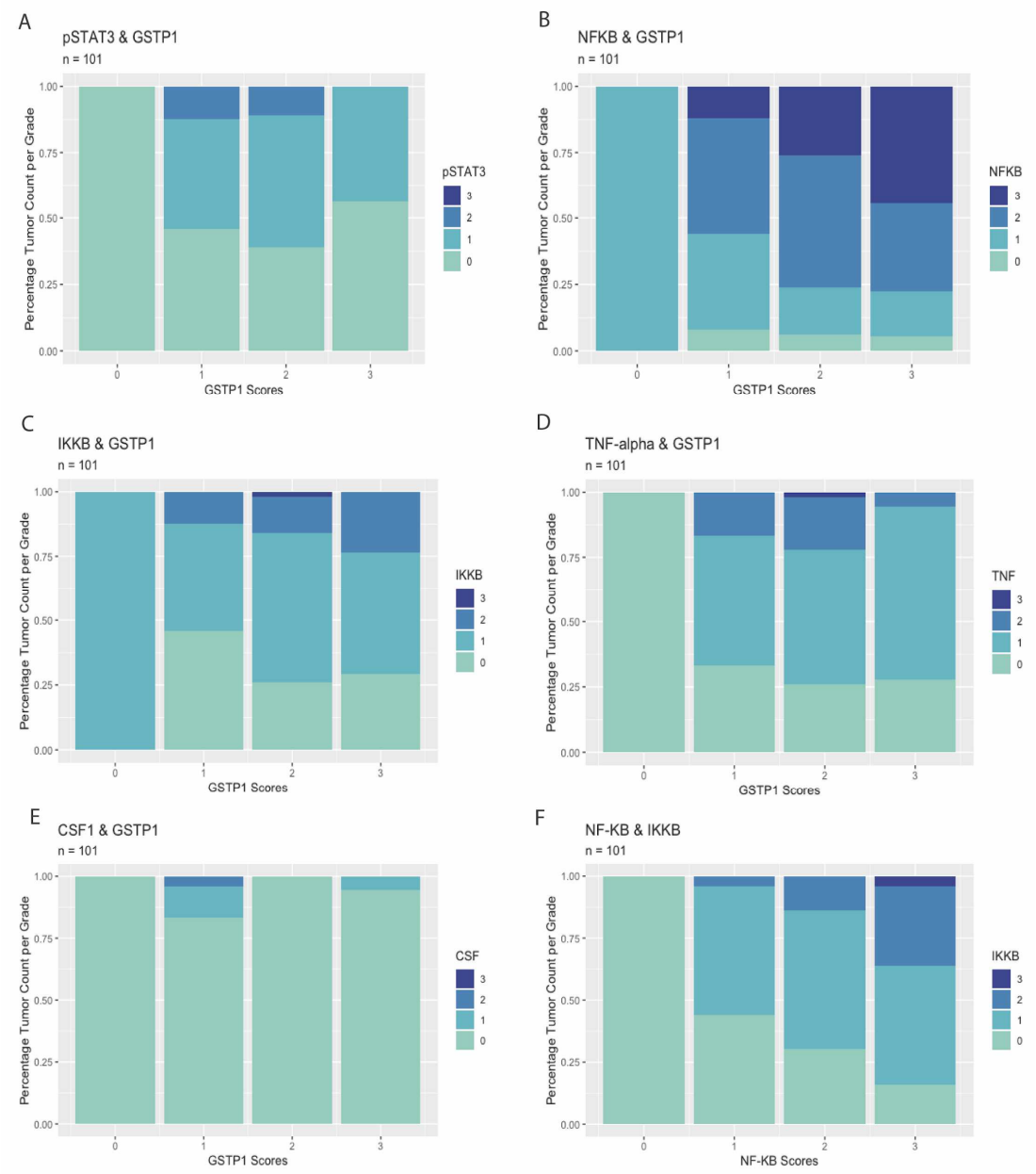

Figure 16: Score Correlation Between GSTP1 Staining and Other Target Genes. A: Staining correlation between pSTAT3 and GSTP1. **B:** Staining correlation between NF-KB and GSTP1. **C:** Staining correlation between IKK and GSTP1. **D:** Staining correlation

between TNF-alpha and GSTP1. **E:** Staining correlation between CSF1 and GSTP1. **F:** Staining correlation between NF-KB and IKKB.

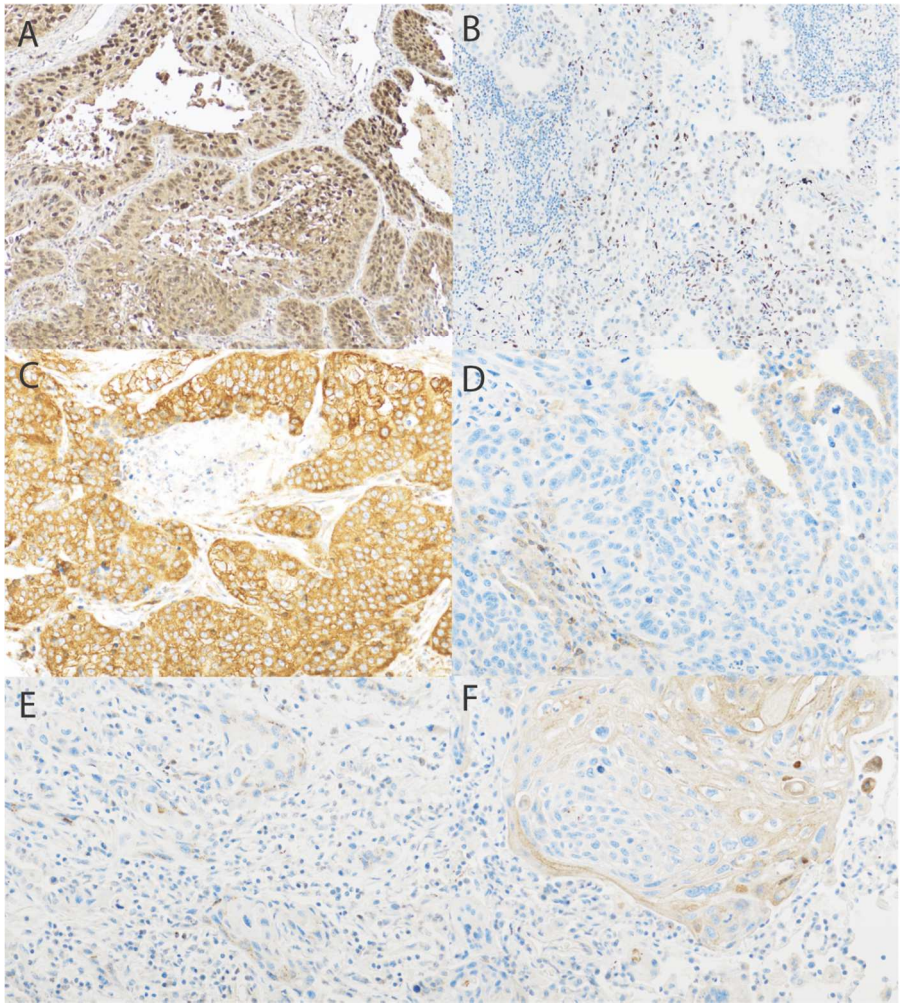

Figure 17: Immunohistochemistry Staining of Target Genes in Human Lung Squamous Cell Carcinoma Tumor. A: Staining of GSTP1 in SCC tumor, score 3; **B:** Staining of pSTAT3 in SCC tumor, score 1. The presence of phosphorylated STAT3 staining suggests that GSTP1 is not directly inhibiting the pathway; **C:** Staining of NF-KB

in SCC tumor, score 3; **D:** Staining of IKKB in SCC tumor, score 0. The non-detection of IKKB in the tumor tissue suggests that IKKB was degraded and NF-KB was released; **E:** Staining of CSF1 in SCC tumor, score 1; **F:** Staining of TNF-alpha in SCC tumor, score 2.

Optimization of DNA and RNA Isolation Methods from Frozen Lung Tissue

In order to optimize concentration of genetic materials being extracted and isolated from the O.C.T embedded frozen mice lung tissue, different tissue sections of different thicknesses were sectioned. We isolated DNA and RNA from 80, 100 and 120µm worth of lung tissue. Following DNA/RNA Kit protocol from QIAGEN, 100µL of buffer EB, we have used ultra-filter water instead, should be used to elute the DNA from the spin column, and 30-50µL of RNase-free water to elute RNA from the RNA spin column. However, when eluting DNA with such a high volume of water, it yielded in a DNA concentration that would not be ideal, when compared to the concentrations that were reached when diluting DNA with lower amounts of water (Table 8). However, results can vary between different samples as not all lung samples are the same size.

The Spectrophotometer Nanodrop also provides values for 260/280 wavelength and 260/230 wavelength. The 260/280 wavelength is a ratio reflective of protein contamination. Ratios between 1.8 and 2.0 are considered pure for DNA, and ratios between 1.8 and 2.1 are considered free of protein contamination for RNA (Koetsier, n.d.). Meanwhile, the 260/230 wavelength is a ratio that measures nucleic acid purity. It is known that ratios below 1.8 alerts for contamination of the nucleic acid in your sample (Koetsier, n.d.).

Results for the 260/280 wavelength ratios were mainly in between the 1.8 and 2.0 range, and interestingly, it was observed that, with a higher amount of tissue being put into the lysis buffer and with a lower volume of elution, results for that 260/280 wavelength were better (Table 9). On the other hand, results for the 260/230 wavelength ratio were much lower than the expected ratio above 1.8. All DNA samples that were measured, regardless of the amount of lung tissue and volume of elution, presented a very low ratio below 1, indicating that the nucleic acid was not pure (Table 10). The 16-week control samples presented the highest nucleic acid purity, with very low deviation in between samples (Table 10). However, this was not an issue when isolating DNA from the frozen tails, which led us to believe that the O.C.T compound in which the lung samples were embedded on could have led to these results.

	Control 722-3	**Control 722-3**	**Control 722-3**
Tissue thickness	100 µm	100 µm	100 µm
Total Volume	150 µl	120 µl	60 µl
DNA Conc.	24 ng/µl	19.5 ng/µl	37.3 ng/ µl

Table 8: **Spectrophotometer Nanodrop Data for DNA Isolation of lung samples at Different Tissue Thickness**. Control lung samples that were cut into the same thickness, however, eluted at three different volumes: 150, 120 and 60 µl. It is observable that samples eluted with higher amount of ultra-pure water showed a lower total DNA concentration, 24 and 19.5 24 ng/µl, when compared to an elution with 60 µl, which led to a total DNA concentration of 37.3 24 ng/µl.

	Control 722-3	**Control 722-3**	**Control 722-3**

Tissue thickness	40 µm	80 µm	100 µm
Total Volume	120 µl	120 µl	60 µl
260/280	2.25	2.06	2.04

Table 9: Spectrophotometer Nanodrop Data for Protein Contamination of DNA Isolation of lung samples. A260/A80 ratios when evaluated at different tissue thickness and volume of elution. It was observed that an increase in tissue amount and a decrease in ultra-pure water for elution led to an improvement of the protein contamination ratio.

	8W NTCU (n = 3)	**12W NTCU** (n = 3)	**16W NTCU** (n = 3)	**8W Control** (n = 2)	**12W Control** (n = 2)	**16W Control** (n = 2)
260/230	0.49 ± 0.22	0.53 ± 0.26	0.41 ± 0.48	0.48 ± 0.33	0.59 ± 0.65	1.7 ± 0.12
Total volume	60 µl	60 µl	60 µl	30 µl	30 µl	30 µl

Table 10: Spectrophotometer Nanodrop Data for Nucleic Acid Purity of DNA Isolation of lung samples. The mean ± sd of A260/230 ratios, evaluated at multiple different scenarios: different tissue thickness, different volume for elution, control samples vs. NTCU treated samples, however, no sample showed a ratio higher than 1.8.

The isolation of genetic material from the tails of the mice had overall great numbers. The concentration of DNA present was high, with some samples reaching somewhere around 450 ng of DNA per µl (Table 11). A260/280 ratios were excellent all around, in which ratios of individual samples varied in between 1.8 and 2.0, implicating that there was no protein contamination (Table 12). Aforementioned, A260/230 ratios were

above 1.8, which indicated that there is no contamination in the nucleic acid within the sample (Table 12).

	8W NTCU (n = 3)	**12W NTCU** (n = 3)	**16W NTCU** (n = 3)	**8W Control** (n = 2)	**12W Control** (n = 2)	**16W Control** (n = 2)
DNA Conc. (ng/μl)	130.63 ± 26.43	185.43 ± 44.91	586.87 ± 286.96	391.8 ± 268.84	281.15 ± 248.5	373.35 ± 12.23

Table 11: **Spectrophotometer Nanodrop Data of DNA Concentration from Tail Samples of Control Mice and NTCU Treated Mice**. The mean ± sd DNA concentration (ng/ μl) for each time point. It is observable that overall mean DNA concentrations were very high.

	8W NTCU (n = 3)	**12W NTCU** (n = 3)	**16W NTCU** (n = 3)	**8W Control** (n = 2)	**12W Control** (n = 2)	**16W Control** (n = 2)
A260/280	1.95 ± 0.03	1.98 ± 0.006	1.93 ± 0.02	1.78 ± 0.23	1.89 ± 0.03	1.97 ± 0.03
A260/230	1.97 ± 0.12	1.96 ± 0.12	2.14 ± 0.1	1.71 ± 0.57	1.97 ± 0.24	2.13 ± 0.007

Table 12: **Spectrophotometer Nanodrop Data of Protein Contamination and Nucleic Acid Purity Ratios in Tail DNA samples.** The mean ± sd of A260/280 ratios and A/260/230 ratios for tail DNA samples. Protein contamination ratios (260/280) and nucleic acid ratios (260/230).

The DNA concentrations results from the nanodrop were not so concerning when compared to RNA quantification results. RNA concentrations were considerably lower

than DNA concentration coming from the lung tissue sections. The RNA from samples were eluted to 60 µL, which for some samples, resulted in decent RNA concentration. It was also observed that the control lung samples had a considerably higher RNA concentration when compared to the NTCU treated samples, which could be attributed to the inflammation levels of the NTCU lungs when under the carcinogenic treatment (Table 13).

Protein contamination and nucleic acid purity levels show worse results than for previous DNA ratios. Most A260/280 ratios are either below 1.8 or above 2.1, which may indicate protein contamination. Furthermore, A260/230 ratios are below 1.8, indicating that the nucleic acid material within the samples are not pure (Table 14). These values can be attributed to the presence of O.C.T compound, however, further analysis was to determine the viability of the sample to be processed and used for sequencing.

	8W NTCU (n = 3)	**12W NTCU** (n = 3)	**16W NTCU** (n = 3)	**8W Control** (n = 2)	**12W Control** (n = 2)	**16W Control** (n = 2)
RNA Conc.	13.03 ± 5.65	12.17 ± 4.98	7.97 ± 2.84	23.8 ± 8.34	15.5 ± 3.39	15.75 ± 12.66

Table 13: **Spectrophotometer Nanodrop Data for RNA Concentration of Control and NTCU Treated RNA Samples.** RNA concentration from NTCU treated lung samples and control lung samples. Values are mean ± standard deviation for each of the samples within a time point (8 weeks, 12 weeks and 16 weeks). All final volumes are 30 µl.

	8W NTCU (n = 3)	**12W NTCU** (n = 3)	**16W NTCU** (n = 3)	**8W Control** (n = 2)	**12W Control** (n = 2)	**16W Control** (n = 2)
A260/280	1.89 ± 0.25	2.15 ± 0.17	2.21 ± 0.19	1.99 ± 0.35	2.18 ± 0.03	2.05 ± 0.08
A260/230	0.16 ± 0.12	0.17 ± 0.16	0.14 ± 0.10	0.59 ± 0.09	0.39 ± 0.02	0.72 ± 0.71

Table 14: **Spectrophotometer Nanodrop Data for Protein Contamination and Nucleic Acid Purity in RNA Samples.** The mean ± sd for A260/280 and A260/230 ratios for all RNA samples, combined by time point for treatment, for both NTCU treated lungs and control lungs. All samples were 150 μm thick and total volume was 60 μl.

The RNA samples went into further analysis of quality. The BioAnalyzer was used in order to identify RNA integrity, also known as the RIN number. The integrity number ranges in scale from 1 to 10, with 10 reflecting a pure RNA material. The RIN number is an analysis of two peaks: the 18S ribosomal RNA around 2000 nucleotides and the 28S ribosomal RNA at around 4000 nucleotides, in which their presence determines the RIN number. A high RIN will have the presence of both peaks, meanwhile, a low RIN will most likely have short peaks for both rRNA subunits and large peaks for degraded RNA (Figure 18A & B). It is important for both subunits of rRNA to be present, seen that rRNA comprises around 80% of total RNA. The presence of these peaks inflicts that the RNA sample has a high integrity, seen that both subunits of rRNA are present. Other peaks found between 25 and 200 nucleotides reflect RNA that have been degraded. The peaks can also be observed in a gel (Figure 19). For most RNA samples analyzed by the BioAnalyzer, RIN numbers were mostly high, ranging in between 7 and 8 (Table 15). Some samples had lower RIN numbers, which could possibly be attributed to either improper handling of the

sample throughout the protocol, contamination, or due to the presence of high amounts of apoptotic cells that contained degraded RNA and inflammation within the lungs. It is important to point out control lung samples mostly had higher RIN numbers than NTCU treated lung samples (Table 15).

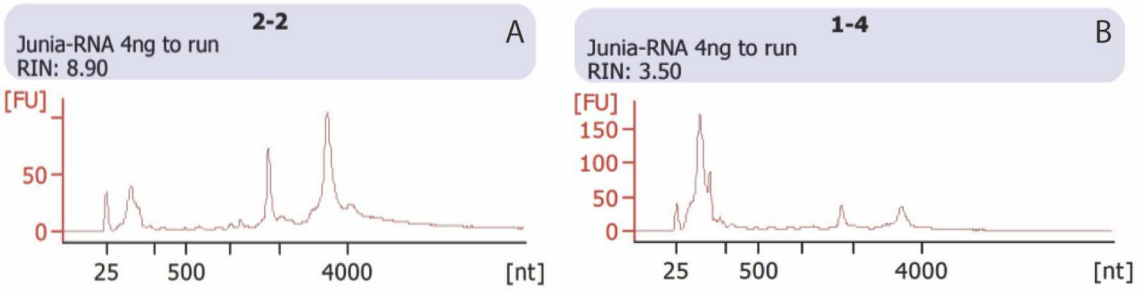

Figure 18: **Agilent BioAnalyzer Data for Qualification of RNA Samples. A:** BioAnalyzer results from an 8-week control lung RNA sample. The high RIN number, 8.90, contains tall peaks for both 18S rRNA, at 2000 nt, and 28S rRNA, at around 4000 nt. Smaller peaks in between 25 and 250 nt correspond to the portion of RNA in the sample that has been degraded, which reflects that little RNA was degraded in this sample; **B:** Bioanalyzer results from a 12-week NTCU treated lung RNA sample. The lowest RIN number, 3.50, results in small peaks around 2000 and 4000 nts, and most peaks are concentrated around 25 and 500 nts, where degraded RNA is measured.

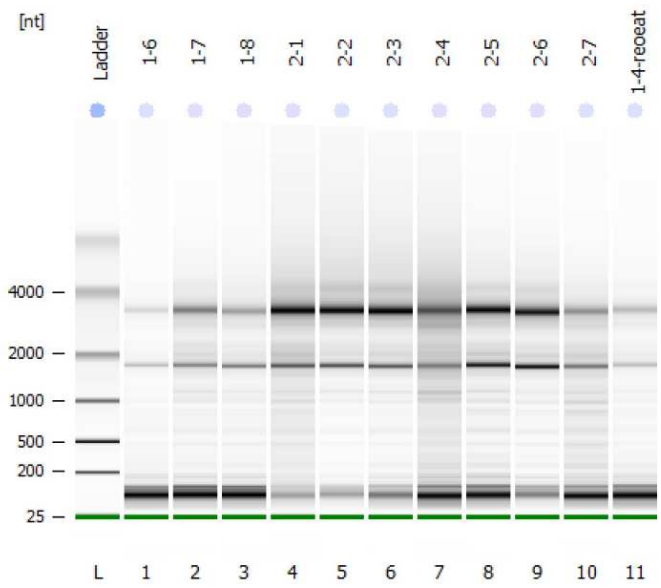

Figure 19: **Gel Electrophoresis Data from Agilent BioAnalyzer for RNA samples** / The gel contains NTCU treated RNA samples, 1-4, 1-6 through 1-8 and 2-7, and control RNA samples, 2-1 through 2-6. In the gel, the darker bands between 25 and 200 nt are degraded RNA material. The band around 2000 nt is the 18S rRNA, and the band approximate to 4000 nt is the 28S rRNA. Control RNA samples show darker bands around rRNA peaks and lighter bands around degraded RNA peaks.

	8W NTCU (n = 3)	**12W NTCU** (n = 3)	**16W NTCU** (n = 3)	**8W Control** (n = 2)	**12W Control** (n = 2)	**16W Control** (n = 2)
RIN	6.75 ± 0.07	5.05 ± 2.19	6.25 ± 0.07	8.25 ± 0.9	6.55 ± 2.62	8.25 ± 0.49

Table 15: Agilent BioAnalyzer Data for RIN Numbers of RNA Samples. RNA Isolation results for the integrity of samples at different time points as mean ± sd. Control RNA samples showed an overall higher RIN numbers when compared to NTCU treated RNA samples.

NTCU Mouse Model & GSTP1 Expression Analysis

Utilizing the methods for isolation of RNA from NTCU mouse tissue, we aimed to determine if the same pattern of expression for GSTP1 and downstream targets observed in cancer cell lines were also valid for NTCU treated mice. For the analysis of the target genes (GSTP1, STAT3, NF-KB, CSF1 and TNF-alpha) and their gene expressions levels in 8-week control (n=2) and 8-week NTCU treated mice (n=2), we have obtained 375μm of lung tissue of both control and NTCU treated mice embedded in OCT. The tissue curls went through RNA isolation protocol, in which QIazol (QIAGEN) and chloroform were used to disrupt the cells as well as to maintain RNA integrity. Both control and NTCU treated RNA samples were subjected to cDNA synthesis, and subsequently qPCR. The TaqMan primers used for this assay were mouse specific primers for GAPDH, GSTP1, STAT3, NF-KB, CSF1 and TNF-alpha.

We were able to observe that there was a significant difference between GAPDH C_T values in mice and SW900 cell lines derived from lung squamous cells carcinoma in humans, 25.04 and 20.59, respectively (Table 16). The GAPDH C_T values for the cell lines were observed to be maintained quite low throughout experiments, even upon treatment. However, it was observed that GAPDH C_T values in mice had a slight decrease in amplification in control mice, as well as NTCU treated mice, when compared to human cell lines (Table 16). Most target genes in mice had significantly higher C_T values when compared to cancer cell lines. GSTP1 did not seem to be upregulated in the control nor the

NTCU treated mice, with C_T values of 28.77 and 31.69, respectively. NF-KB presented a significant downregulation upon NTCU treatment (Table 16). Whereas, TNF-alpha expression did not change throughout. The remaining genes, CSF1 and STAT3, showed slight decrease in C_T values.

Two different mice were analyzed in this experiment, and even though there were variances between one mice lung to another, both mice presented low expression of GSTP1, in both control and NTCU treated lung (Figure 20A). There were some differences in between mice regarding gene expression, therefore, explaining slightly higher standard deviation values (Table 16). Both mice presented high relative expression values for TNF-alpha, however, TNF-alpha had a low amplification curve and remained fairly constant from control to NTCU treated mice, and therefore, the large relative expression value can be attributed to the change in C_T value of GAPDH from control to NTCU treated lung.

	AALE Cells	**H2170 cells**	**SW900 cells**	**8W Control Mice**	**8W NTCU Mice**
GAPDH	20.75 ± 0.20	21.34 ± 2.00	20.59 ± 1.29	25.04 ± 0.13	28.57 ± 1.14
GSTP1	19.92 ± 0.04	20.36 ± 2.89	21.43 ± 1.39	28.77 ± 0.09	33.83 ± 2.49
STAT3	24.98 ± 0.01	25.19 ± 0.48	24.79 ± 0.59	28.98 ± 0.03	31.46 ± 1.84
NFKB	23.91 ± 0.06	25.04 ± 0.45	25.48 ± 0.93	30.46 ± 0.02	34.31 ± 1.33
CSF1	28.93 ± 0.03	25.05 ± 0.87	27.29 ± 0.56	30.21 ± 0.13	32.73 ± 0.84
TNF-alpha	32.59 ± 0.30	33.89 ± 0.65	36.83 ± 0.13	35.59 ± 0.39	36.15 ± 0.65

Table 16: Comparison of C_T values Of Target Genes in Cell Lines of Interest, 8-Week Control Mice and 8-Week NTCU Treated Mice. The mean±sd of C_T values across cell lines of interest, AALE, H2170 and SW900, and mice lung tissue at 8 weeks. It was observable that GSTP1 was significantly downregulated in mice when compared to cell lines.

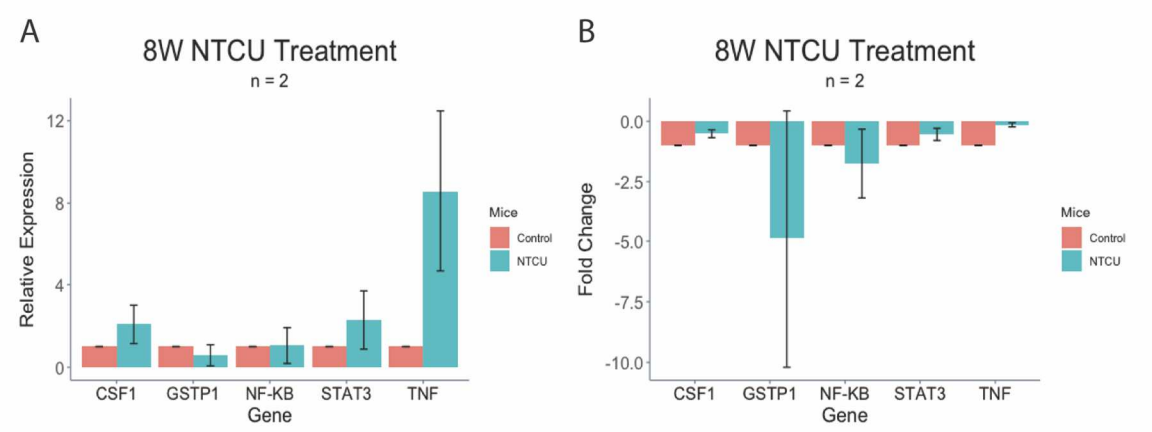

Figure 20: Gene Expression of Target Genes in NTCU Mouse Model. A: Relative expression is demonstrated as mean ± sd of $2^{-\Delta\Delta CT}$ values, normalized for GAPDH. GSTP1, and NF-KB to some extent, were fairly lower expressed in NTCU treated mice when compared to control mice. **B:** Mean ± sd of fold change in target genes in mice lung samples. There was an increase in the down-regulation of GSTP1 and NF-KB, when compared to other genes, on NTCU treated mice, however, all genes presented diminished expression in NTCU mice.

Hematoxylin & Eosin Stained Control and NTCU Treated Lung Samples

To further associate the information from the PCR studies to the presence of histology we examined the lesions in the tissue from which the RNA was isolated. Tissue

sections, for both control and NTCU treated lung samples, were placed in the glass slide during the collection of tissue for DNA and RNA Isolation. An 8 μm tissue section was obtained before the tissue curls were acquainted for genetic material extraction, and a second section was obtained once the curls had already been collected. The tissue sections were placed in the same microscope slide and further stained with Haematoxylin & Eosin (H&E) for the visualization of both the cytoplasm and nuclei at different spectrum of colors, seen that each stain is specific to each cellular component. Furthermore, the tissue sections were obtained before and after tissue curls collection to ensure that the genetic material being isolated for those samples had in fact a PML histology.

When analyzing the slides under the microscope, in the control samples at 8 weeks, first time period in which the lungs were harvested, fairly normal epithelium was observed within bronchi opening, however, these observances were local and the air sacs remained intact and clean from any inflammation (Figure 21A). Control samples at 12 weeks showed some increase of inflammation within the lung (Figure 21E), along with some atypia (Figure 21B). Furthermore, control lung samples at 16 weeks, which was the last time point in which lungs were harvested, showed some degree of hyperplasia and some potential fibrosis, but these samples are heterogeneous, and therefore, some areas of the lung seem to show different histological grades, such as some areas of flat atypia (Figure 21C and 21D). There is a visible increase of inflammatory cells infiltration in the 16-week NTCU samples (Figure 22G and 22H). No specific staining, such as immunofluorescence or immunohistochemistry, was done in order to identify if the cells infiltrating the air sacs of these lungs were indeed immune cells, however, as the NTCU treatment is topically applied

and more recently suggested may induce disease through inhalation, it is possible that control mice cohabiting in the same racks as treated mice could lead to some low grade changes of the epithelium as seen in the control mice, as well as inflammatory responses.

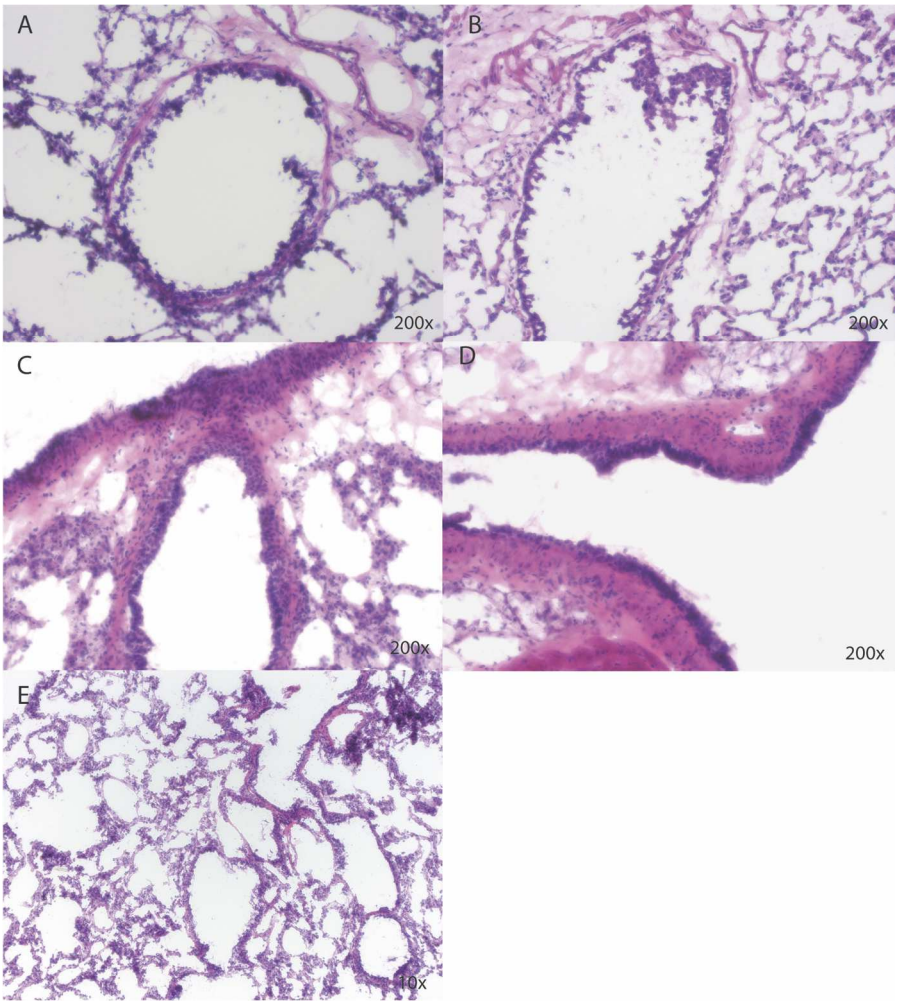

Figure 21: Histology of Control Lung Samples at 8, 12 and 16 weeks, stained with H&E. A: 8-week control lung, showing squamous metaplasia ; **B:** 12-week control lung presenting squamous metaplasia; **C:** 16- week control lung, with bronchial hyperplasia. Depicts unorganized cellular architecture within bronchi epithelium; **D:** 16-week control lung, focusing on lining of airways, with bronchial hyperplasia. Cells are highly

unorganized and basal layer of cells are no longer visible; **E:** 12-week control lung, at lower magnification, showing increased presence of inflammatory cells within air sacs.

The histology of the tissue sections coming from NTCU treated mice at 8, 12 and 16 weeks had more apparent changes to epithelium, as well as air sacs, when compared to control mice. The lung of mice treated with NTCU presented certain abnormalities to the linings of the bronchi. The normal pseudostratified epithelium observed in healthy lungs has gone through observable changes, as atypical cells start to transform the pseudostratified columnar epithelium into squamous epithelium, in which the presence of disorganized cells is visible. Furthermore, the histology of the 8-week, post a single NTCU treatment, lung presents low-grade dysplasia, and the nuclei to cytoplasm ratio seems to have decreased (Figure 22B). The presence of horizontally lined atypical cells is also seen (Figure 22A). When analyzing the second time point of the treatment, mice lungs at 12 weeks post NTCU have an increase in inflammation, when compared to 8-week post NTCU treated lungs, which is observable by the increased presence of cells within the lung tissue (Figure 22C and 22D). Moreover, it was also visible that the epithelial lining of the bronchi presented low-grade dysplasia, similarly to what was observed at 8 weeks. The lungs harvested from mice at 16 weeks of NTCU treatment were histologically graded as low-grade dysplasia, in which the basal layer of the epithelium is still intact, however, upper layers are disorganized and have matured (Figure 22E). One lung sample from 16 week mice was graded as flat atypia (Figure 22F), in which the epithelium is somewhat normal, composed of one cell layer, however, it shows enlarged nuclei.

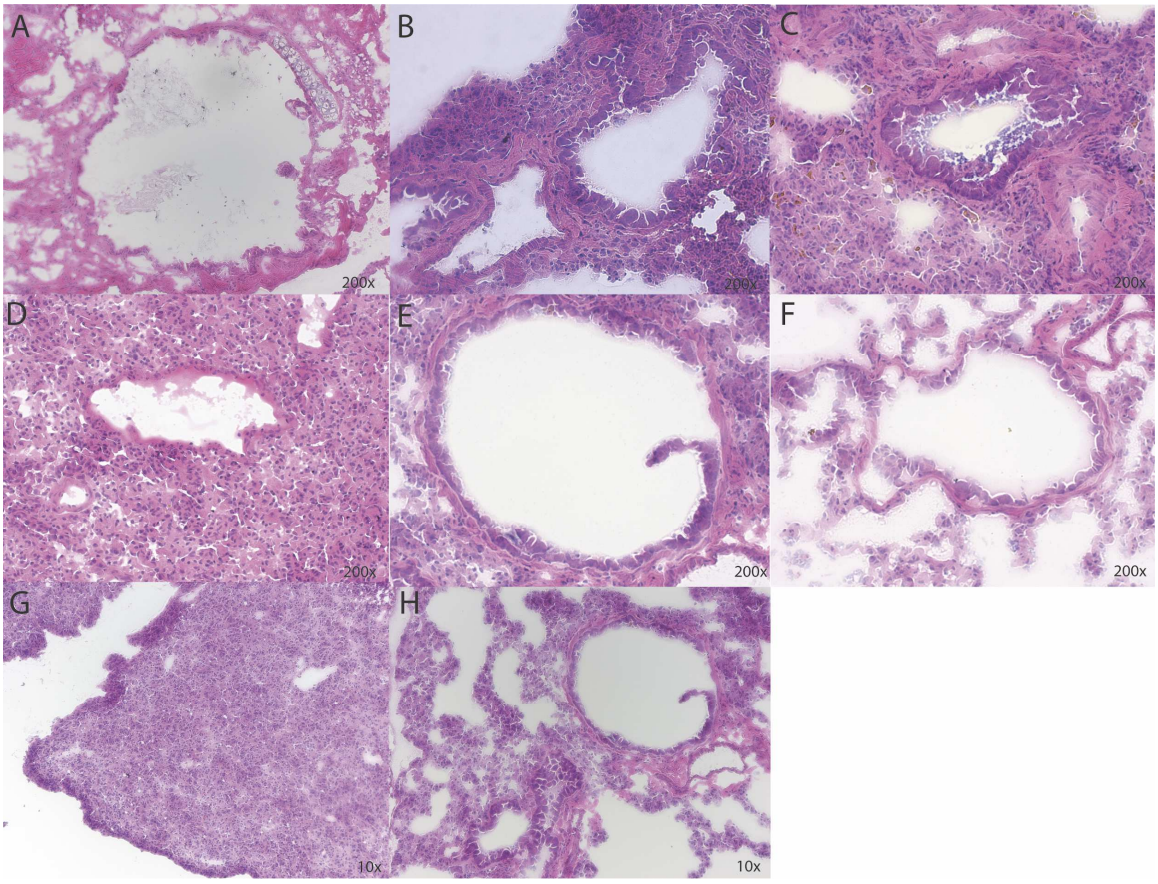

Figure 22: Histology of NTCU Treated Lung Samples at 8, 12 and 16 weeks, stained with H&E. A: 8-week NTCU treated lung graded as low-grade dysplasia; **B**: 8-week NTCU treated lung from mice, presenting low-grade dysplasia lesions; **C:** 12-week NTCU, graded as low-grade dysplasia, with significant increase of inflammatory cells; **D:** 12-week NTCU, graded as low-grade dysplasia; **E:** 16-week NTCU with low grade dysplasia; **F:** 16-week NTCU presenting flat-atypia within the lining of bronchi, with enlarged nuclei; **G:** 16-week NTCU, at lower magnification, showing increase inflammation on lung tissue section; **H:** 16-week NTCU, at low magnification, showing inflammatory cells within air sacs.

For the purposes of isolating genetic material from frozen samples, embedding the harvested lungs in O.C.T is necessary. However, preparing slides from O.C.T embedded samples could be prejudicial to the histological analysis of the sample. Frozen samples when embedded in O.C.T are sensitive and their manageability is limited. Therefore, in order to properly analyze the histology of the 8, 12 and 16-week NTCU treated mice and control mice, we have prepared FFPE slides from the same mice, however, with a different portion of the same harvested lung used for DNA/RNA Isolation. It is noticeable that in FFPE fixed lung samples, when sliced and placed on slides, lung architecture is maintained (Figure 23A, 23C and 23D). Whereas, O.C.T embedded samples mostly have compromised architecture of important features of the lung, such as the bronchi (Figure 23B and 23F). Tissue overlapping is also very common while placing the tissue slice onto the glass slide (Figure 23D), which causes a misleading histological figure of the sample.

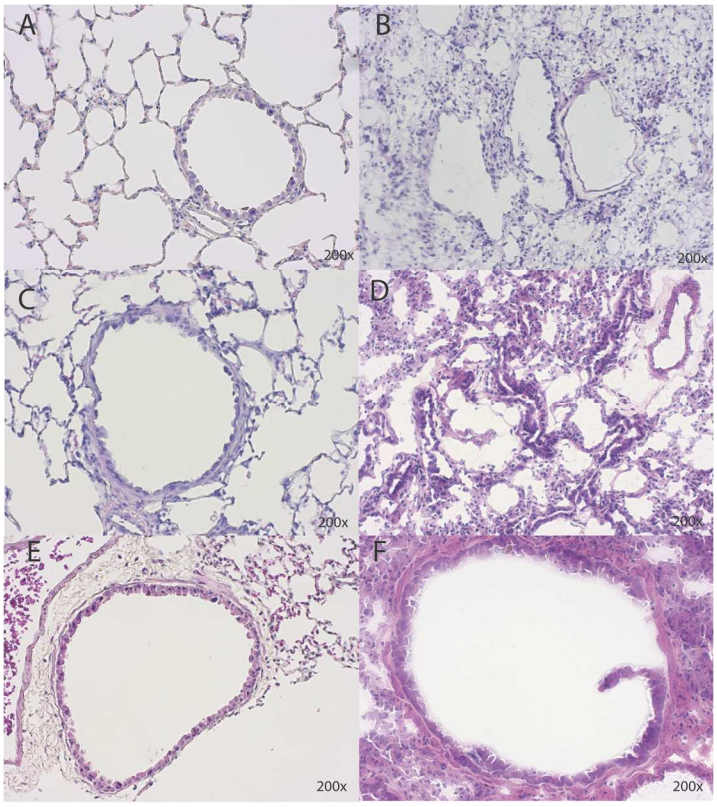

Figure 23: Comparison between FFPE embedded lungs and Frozen O.C.T embedded lungs. A: 8-Week NTCU treated FFPE fixed lung; **B:** 8-week NTCU treated frozen lung. The overall architecture of bronchi is highly disrupted. Inflammatory cells infiltrate is significantly higher than in FFPE slides; **C:** 12-week NTCU treated FFPE fixed lung; **D:** 12-week NTCU treated frozen lung. Overlapping area of tissue due to sensitivity of O.C.T embedded tissue; **E:** 16-week NTCU treated FFPE fixed lung. Architecture of the lung, more specifically bronchi, is intact; **F:** 16-week NTCU treated frozen lung. The figure depicts a bronchus in which had most of its structure maintained, however, it is possible to observe a flap of epithelial tissue loose.

The histology of the lung from 8-week control mice shows no inflammation signs, such as presence of inflammatory cells within air sacs, and the lining of the epithelium appears normal like. When analyzing the correlation of the histology with the expression of target genes (GSTP1, STAT3, NF-KB, CSF1 and TNF-alpha), which were all downregulated, suggests that the low expression of these genes may be a result of their translation not being activated seen that no inflammation is observed within the lungs of control mice. However, in the 8-week post NTCU treated lungs, the histology of the lining of the bronchi presents a low-grade dysplasia, but inflammatory cells are not really observed within the air sacs. In these mice, the target genes were further downregulated than the control mice, and this may suggest that the correlation between histology and gene expression is that with further worsening of these lesions, these immune pathways may be getting inhibited. However, in mice, GSTP1 does not appear to be the factor inhibiting STAT3 and NF-KB pathways from translating their downstream genes and further releasing cytokines and chemokines for the development of an immune response.

DISCUSSION

The diagnosis of early stage lung cancer is an unmet need, along with identification of therapeutic targets. Extensive work has been done on advanced stages of lung cancer, as that is typically when tumors are distinctive from the normal epithelium of the airways and present significant abnormal changes to histology. However, prognosis of advanced stages are low, and along with that arises the necessity for identifying transcriptional, genetic and microenvironmental changes in premalignant lesions of LUSC that will allow for in depth understanding of these lesions. For this project, we have aimed to first validate computational analyses that suggested GSTP1 to be anticorrelated with immunoregulatory pathways such as STAT3 and NF-KB. We have analyzed the hypothesis of GSTP1 being a possible inhibitor of immune responses through these pathways by the use of two models: cancer cell lines and tissue microarrays from stage I and stage II LUSC tumors. Along with baseline expression of target genes in these models, we have further the studies by interfering with GSTP1 expression and observing the effects on STAT3 and NF-KB pathways. Our second aim focused on bridging the first aim with a mouse model known as NTCU. We intended to observe the relationship of GSTP1 and STAT3 and NF-KB pathways in lung tissue from NTCU treated mice, as well as further validate the model as viable for the study of PMLs in humans.

The first aim of this study consisted of analyzing the relationship between GSTP1 and STAT3 and NF-KB pathways in cancer cell lines. As previously mentioned, GSTP1 has been identified to be anti-correlated with the expression of STAT3, NF-KB, CSF1 and

TNF-alpha. We were able to determine that in cancer cell lines such as SW900 and H2170, which are cancer cell lines originated from lung squamous cell carcinomas, as well as immortalized cell lines presented a high expression of GSTP1, meanwhile, the downstream targets were downregulated, which was in concordance with computational analysis studies. We, then, aimed to knock down GSTP1 through RNA interference protocol, to observe if there was an increase in the expression of STAT3 and NF-KB pathways. Upon transfecting the SW900 cells with GSTP1 siRNA, to promote the knockout of the gene, it was observed that GSTP1 presented a significant decrease with a 25 fold-change, however, the proposed impact on the downstream immune targets did not increase but rather presented slight decrease in fold-change, which was not expected at first. We believe this could be attributed to the fact that GSTP1 is not modulating these pathways in this context potentially because there is no immune stimulation or carcinogen/tobacco exposure that could cause these genes to become upregulated.

In effort to test the effects of an immune stimulation on STAT3 and NF-KB pathways upon GSTP1 knockdown, we treated the cells with interleukins, to observe if inhibiting GSTP1 and stimulating an immune response would lead to an increase of the STAT3 and NF-KB pathways. Upon stimulation of immune responses with IL-13 and IL-8 for 24 and 48 hours, there was again an observance of a slight decrease in fold-change of STAT3, NF-KB, CSF1 and TNF-alpha. Moreover, the interleukin stimulation for a 4-hour time period was thought to capture the interleukins peak of activity, however, for IL-13 and IL-8, there was no upregulation of any of the downstream targets. Previous work suggests TNF-alpha can act as one of many ligands that can bind to a cell surface receptor

activating the canonical NF-KB and STAT3 pathways, and has been suggested that it may also be up-regulated in some cancers (Lu, H., Lei, X. & Zhang, Q., 2015, Rébé et al., 2013). However, in the PCGA data and our qPCR results have shown that TNF-alpha expression levels are very low, even upon interleukin stimulation. Low expression of TNF-alpha could potentially implicate why we didn't see the activation of NF-KB and STAT3 pathways, explaining the downregulation of these genes at mRNA levels. It is not well understood how IL-13 and IL-8 stimulation may drive immunoregulatory pathways in our model, however our results suggest receptors for IL-13 and IL-8 may not be expressed in the cancer cell line used in the study or they could have potentially lost its ability to be stimulated by IL-13 and IL-8.

Following the inability for the knockdown of GSTP1 and stimulation of the cells with interleukins to trigger an immune response activating the STAT3 and NF-KB pathways, we were intrigued to observe the effects of H_2O_2 and CSC on GSTP1 in SW900 cancer cells. Upon a 4 hour treatment of these cells with H_2O_2, the only target gene that was observed to be affected by the induction of oxidative stress was TNF-alpha, which was up-regulated in comparison to control cells. The increase in relative expression of TNF-alpha suggests that an increase in production of reactive oxygen species (ROS) triggered production of TNF-alpha as a pro-inflammatory cytokine and immune response, consequently leading to a broader inflammatory response. The increase in production, and therefore, expression of TNF-alpha suggest that the NF-KB pathway was activated upon the presence of H_2O_2. Previous studies on human lung epithelial cells have observed that there is a relationship between TNF-alpha and H_2O_2, in which enhanced exposure of cells

to TNF-alpha, increased H_2O_2 production in human lung epithelial cells (Babbar & Casero, 2006). Interestingly, even upon up-regulation of TNF-alpha, the relative expression of GSTP1 did not change, suggesting that GSTP1 may not regulate the NF-KB pathway directly as we had hypothesized. Furthermore, the STAT3 pathway seems to not have been activated by the presence of H_2O_2 as the relative expression of both STAT3 and CSF1 did not appear to have any changes between treated and untreated cells.

The addition of cigarette smoke condensate to SW900 cells was expected to cause an increase in the oxidative stress of the lung epithelial cells, seen that CSC is composed of chemical particles that are filtered from cigarette smoke. As a result of an increase of oxidative stress of cells, it is suggestive that GSTP1 expression, as being involved in the xenobiotic metabolism, would be maintained at high levels. However, upon 4 hour treatment of cells with CSC, we were only able to observe a decrease in genes such as CSF1 and TNF-alpha, suggesting that the transcription of these genes were not activated upon CSC treatment and that, therefore, inflammatory/immune responses were also not activated. However, an increase in treatment time with CSC may be better effective to analyze the response of STAT3 and NF-KB pathways to smoke condensate, allowing for longer exposure time.

Furthermore, using qPCR for measuring the expression of STAT3 and NF-KB was not sufficient to determine if the knock down of GSTP1 did allow these pathways to be activated, seeing that CSF1 and TNF-alpha gene expressions were maintained very low throughout these experiments. Further experiments analyzing the phosphorylation of

STAT3 and IkBa through Western Blot, or an assay determining localization of these proteins within the nucleus, are necessary to determine their activation.

In effort to look at the expression of GSTP1 and proposed targets that I perturbed in cancer cell lines, we stained squamous cell carcinoma (LUSC) tumor tissue from human samples with antibodies specific to our target genes. The results of the immunohistochemistry suggest that, similarly to what we have observed in cancer cell lines and the pre-cancer data, GSTP1 is upregulated in LUSC tumor tissues, which yields in high intensity of antibody staining. As GSTP1 is upregulated, in most tumor tissue samples we have stained, pSTAT3 is downregulated in comparison to GSTP1. The lower expression of pSTAT3 suggests that GSTP1 is a possible inhibitor of the phosphorylation of STAT3. As a result, the expression of CSF1 was mostly negative throughout the LUSC tumors stained with CSF1 antibody, however, couple tumor tissues had limited staining. The cytosolic expression of NF-KB in the IHC staining is high across most LUSC tumors we have analyzed, suggesting that NF-KB has not been translocated to the nucleus, and therefore, it is not capable of transcriptionally activating its downstream targets. The correlation of high expression of GSTP1 and high expression of cytosolic NF-KB in ICH contributes to the suggestion that upregulation of GSTP1 can inhibit activation of NF-KB. In concordance with high expression of cytosolic NF-KB, is the low expression of TNF-alpha observed across all tumor samples, which is downstream target of NF-KB upon NF-KB translocation to the nucleus. Furthermore, expression of IKKB was also analyzed in order to ensure that NF-KB was activated and translocated to the nucleus for transcription

of its downstream targets. We have observed that increased expression of cytosolic NF-KB is correlated with decreased expression and intensity of IKKB antibody on LUSC tumors.

In the second aim, we wished to determine whether NTCU was a viable mouse model for the study of premalignant lesions of LUSC in humans. Therefore, we have optimized DNA and RNA isolation methods for frozen lung tissue harvested from mice at 8, 12 and 16 weeks of treatment, as well as control mice. We were able to establish that thicker curls of tissue, 150 μm, and low eluate volume, 30 to 60 μl, yielded better results when analyzed with the Spectrophotometer Nanodrop. The DNA samples were observed to be simpler to handle, when compared to RNA samples, as the DNA concentrations were very constant throughout. However, samples that were kept in O.C.T compound for longer, once the curls were cut and the DNA was extracted, showed a slight worsening in the protein contamination and nucleic acid purity ratios. The challenges attributed to RNA isolations yielding in low concentrations could be due to the RNA half-life, which is extremely short once has been extracted from cells and tissues, causing the RNA to become unstable (Tan & Yiap, 2009), and therefore, affecting its quality. However, even though the Nanodrop results showed low concentrations for RNA, the BioAnalyzer, which does a more complete analysis of the sample, resulted in most of the RNA samples having a high integrity, even if their concentrations were not as high. Therefore, knowing the RIN numbers for the samples was important to determine if those samples could be used for transcriptome sequencing.

Furthermore, using material isolated for RNA studies from 8-week control and 8-week NTCU treated lung mice, we were able to synthesize cDNA and run qPCR to observe

if the expression of our genes of interest, GSTP1, STAT3, NF-KB, CSF1 and TNF-alpha, followed the same pattern of expression of that observed in cancer cell lines, SW900 and H2170, and immortalized cell lines, AALE, at baseline. We observed that, by comparing 2 different lung samples from both 8-week control mice and 8-week NTCU treated mice, that GSTP1 was significantly downregulated upon NTCU treatment, however, its expression was observed to be low in the control mice as well. Both NTCU treated mice and control lung samples had lower expression of considerate downregulation of all target genes when compared to the expression of these genes at baseline for cancer cell lines (SW900 and H2170) and immortalized cell line (AALE). The GAPDH expression levels on mice were also observed to drop upon NTCU treatment, which is usually not observed in cell lines. The difference in expression between GAPDH in control and treated mice could compromise the comparison between the other target genes, as GAPDH is being used as a reference gene. However, further studies may include the use of a few different reference genes to run qPCR, allowing for a correct validity of the reference genes as a constant variable, which will ultimately allow for correct biological assumptions to be made regarding qPCR results for the mice lung samples. Furthermore, genes such as CSF1 and STAT3 were mostly observed to be maintained with similar expressions between control and NTCU mice. Furthermore, GSTP1 does not seem to be an inhibitor of STAT3 and NF-KB pathways, seen that in the control mice all genes present low expression. The expression of target genes in NTCU treated mice, as well as control, did not seem to follow the same pattern of expression and correlation found in cancer cell lines and immortalized cell lines. As previously stated, the significant fluctuation of GAPDH expression levels

between control and treated mice may have become a limitation to this study, not allowing an accurate comparison between mice expression levels of our target genes and the cell lines that were used in previous presented experiments. To further the analysis of GSTP1 and downstream targets in mice lungs, it would be beneficial to also analyze lesions from 12-week and 16-week NTCU treated mice, as these mice present higher grades lesions, which are more histologically similar to PMLs observed in human samples.

To also explore the connection between the histology of the lesions and the expression of target genes, the tissue sections coming from the same lung samples in which DNA and RNA isolation was done, were stained with H&E and observed a considerate correlation between RNA concentration and grade of lesion of the histology of the specific sample. Tissue samples that presented a higher-grade lesion with higher inflammation of the lungs, such as few samples from NTCU treated mice at 12 and 16 weeks, also presented a smaller concentration of RNA, even though the same amount of tissue was being collected as the other time points. However, some samples from 12 and 16-week NTCU treated lungs do present a higher concentration of RNA than 8-week NTCU treated mice, and therefore, it is important to take into consideration that not all lung samples were the equal in size. In conclusion, there was a concordance between the histology of the lung sample, its time point of treatment and DNA/RNA concentration being extracted. Furthermore, the correlation between histology and expression of target genes suggests that as lesions progress from normal like epithelium to low grade dysplasia that immunoregulatory pathways, STAT3 and NF-KB seem to be even further downregulated, which leads to very little expression of CSF1 and TNF-alpha in the low grade lesions.

Further studies are ongoing to sequence samples and determine if mutational signatures in NTCU mouse models are similar to those observed in humans. However, we were able to observe that the histology of the premalignant lesions in mice treated with NTCU resembled those observed in PMLs of humans.

In order to determine the role GSTP1 plays in the progression or regression of premalignant lesions of LUSC, more experiments are required. For example, further studies inducing oxidative stress, of cigarette smoke condensate and H_2O_2 for longer treatment periods. Since the interleukin stimulation, at 48, 24 and 4 hours, had no effect on CSF1 and TNF-alpha, the use of a gene as a measure for cell stress would validate that the GSTP1 knockdown study protocols were working properly, and that the suggested inhibitor of STAT3 and NF-KB pathway may not be GSTP1. Moreover, protein studies, such as western blot, maybe need to be assessed to determine if GSTP1 can modulate immune changes. Determining the presence post translational modifications of genes such as phosphorylation of STAT3 and IkBa, will enable insight into the interaction between GSTP1 and these pathways, by in which the presence of phosphorylated STAT3 and IkBa would suggest that GSTP1 may not be regulating these pathways.

The relationship between GSTP1 and the STAT3 and NF-KB pathways and GSTP1 ability to regulate the activation of these genes remains inconclusive. Our studies have suggested that this relationship may not be as straightforward as expected, and that other components may come into play along with GSTP1 and accounting for the downregulation of the immune pathways. However, in our study, we were able to validate baseline expression of our target genes as they are in concordance with computational analyses from

our group. Assays that intervened with GSTP1 expression presented in this thesis have limitations, and as suggested, additional studies could give additional information at the protein level in which qPCR does not. Furthermore, results from the IHC staining on early stages of LUSC tumor samples have elucidated what our experiments on cancer cell lines have been showing, as that GSTP1 may regulate STAT3 and NF-KB pathways to some extent. Further analysis of our target genes with different assays for identification of protein expression and localization could really benefit the purpose of this study, allowing for more accurate conclusion of how GSTP1 regulates the activation of these immuno-regulatory pathways.

REFERENCES

Afridi, F., Ferreira, B. P., & Geigel, E. (n.d.). *A Rare Case Of Papillary Endobronchial Squamous Cell Carcinoma*. 2.

Allocati, N., Masulli, M., Di Ilio, C., & Federici, L. (2018). Glutathione transferases: Substrates, inhibitors and pro-drugs in cancer and neurodegenerative diseases. *Oncogenesis*, *7*(1), 8. https://doi.org/10.1038/s41389-017-0025-3

Babbar, N., & Casero, R. A. (2006). Tumor Necrosis Factor-α Increases Reactive Oxygen Species by Inducing Spermine Oxidase in Human Lung Epithelial Cells: A Potential Mechanism for Inflammation-Induced Carcinogenesis. *Cancer Research*, *66*(23), 11125–11130. https://doi.org/10.1158/0008-5472.CAN-06-3174

Beane, J. E., Mazzilli, S. A., Campbell, J. D., Duclos, G., Krysan, K., Moy, C., Perdomo, C., Schaffer, M., Liu, G., Zhang, S., Liu, H., Vick, J., Dhillon, S. S., Platero, S. J., Dubinett, S. M., Stevenson, C., Reid, M. E., Lenburg, M. E., & Spira, A. E. (2019). Molecular subtyping reveals immune alterations associated with progression of bronchial premalignant lesions. *Nature Communications*, *10*(1), 1–13. https://doi.org/10.1038/s41467-019-09834-2

Campbell, J. D., Alexandrov, A., Kim, J., Wala, J., Berger, A. H., Pedamallu, C. S., Shukla, S. A., Guo, G., Brooks, A. N., Murray, B. A., Imielinski, M., Hu, X., Ling, S., Akbani, R., Rosenberg, M., Cibulskis, C., Ramachandran, A., Collisson, E. A., Kwiatkowski, D. J., … Meyerson, M. (2016). Distinct patterns of somatic genome alterations in lung adenocarcinomas and squamous cell carcinomas. *Nature Genetics*, *48*(6), 607–616. https://doi.org/10.1038/ng.3564

Campbell, J. D., Mazzilli, S. A., Reid, M. E., Dhillon, S. S., Platero, S., Beane, J., & Spira, A. E. (2016). The Case for a Pre-Cancer Genome Atlas (PCGA). *Cancer Prevention Research (Philadelphia, Pa.)*, *9*(2), 119–124. https://doi.org/10.1158/1940-6207.CAPR-16-0024

Cancer Genome Atlas Research Network. (2012). Comprehensive genomic characterization of squamous cell lung cancers. *Nature*, *489*(7417), 519–525. https://doi.org/10.1038/nature11404

Carter, S. L., Eklund, A. C., Kohane, I. S., Harris, L. N., & Szallasi, Z. (2006). A signature of chromosomal instability inferred from gene expression profiles predicts clinical outcome in multiple human cancers. *Nature Genetics*, *38*(9), 1043–1048. https://doi.org/10.1038/ng1861

Cruse, J. M., Lewis, R. E., & Wang, H. (Eds.). (2004). 7—ANTIGEN PRESENTATION. In *Immunology Guidebook* (pp. 267–276). Academic Press. https://doi.org/10.1016/B978-012198382-6/50031-5

Chudgar, N. P., Bucciarelli, P. R., Jeffries, E. M., Rizk, N. P., Park, B. J., Adusumilli, P. S., & Jones, D. R. (2015). Results of the National Lung Cancer Screening Trial: Where Are We Now? *Thoracic Surgery Clinics*, *25*(2), 145–153. https://doi.org/10.1016/j.thorsurg.2014.11.002

de Groot, P. M., Wu, C. C., Carter, B. W., & Munden, R. F. (2018). The epidemiology of lung cancer. *Translational Lung Cancer Research*, *7*(3), 220–233. https://doi.org/10.21037/tlcr.2018.05.06

Duma, N., Santana-Davila, R., & Molina, J. R. (2019). Non–Small Cell Lung Cancer: Epidemiology, Screening, Diagnosis, and Treatment. *Mayo Clinic Proceedings*, *94*(8), 1623–1640. https://doi.org/10.1016/j.mayocp.2019.01.013

Edwards, C., & Carlile, A. (1985). Clear cell carcinoma of the lung. *Journal of Clinical Pathology*, *38*(8), 880–885.

Ghosh, M., Dwyer-Nield, L. D., Kwon, J. B., Barthel, L., Janssen, W. J., Merrick, D. T., & Keith, R. L. (2015). Tracheal Dysplasia Precedes Bronchial Dysplasia in Mouse Model of N-Nitroso Trischloroethylurea Induced Squamous Cell Lung Cancer. *PLoS ONE*, *10*(4). https://doi.org/10.1371/journal.pone.0122823

Gu, J., Hua, F., Mei, C., Zheng, D., Wang, G., & Zhou, Q. (2014). HapMap-based study on the association between MPO and GSTP1 gene polymorphisms and lung cancer susceptibility in Chinese Han population. *Acta Pharmacologica Sinica*, *35*(5), 636–644. https://doi.org/10.1038/aps.2014.11

Hillmer, E. J., Zhang, H., Li, H. S., & Watowich, S. S. (2016). STAT3 signaling in immunity. *Cytokine & Growth Factor Reviews*, *31*, 1–15. https://doi.org/10.1016/j.cytogfr.2016.05.001

Idriss, H. T., & Naismith, J. H. (2000). TNF alpha and the TNF receptor superfamily: Structure-function relationship(s). *Microscopy Research and Technique*, *50*(3), 184–195. https://doi.org/10.1002/1097-0029(20000801)50:3<184::AID-JEMT2>3.0.CO;2-H

Inoue□Choi, M., Hartge, P., Liao, L. M., Caporaso, N., & Freedman, N. D. (2018). Association between long-term low-intensity cigarette smoking and incidence of smoking-related cancer in the national institutes of health-AARP cohort.

International Journal of Cancer, *142*(2), 271–280. https://doi.org/10.1002/ijc.31059

Ishizumi, T., McWilliams, A., MacAulay, C., Gazdar, A., & Lam, S. (2010). Natural history of bronchial preinvasive lesions. *Cancer Metastasis Reviews*, *29*(1), 5–14. https://doi.org/10.1007/s10555-010-9214-7

Ivashkiv, L. B., & Donlin, L. T. (2014). Regulation of type I interferon responses. *Nature Reviews. Immunology*, *14*(1), 36–49. https://doi.org/10.1038/nri3581

Jain, A., & Pasare, C. (2017). Innate control of adaptive immunity: Beyond the three-signal paradigm. *Journal of Immunology (Baltimore, Md. : 1950)*, *198*(10), 3791–3800. https://doi.org/10.4049/jimmunol.1602000

Janjigian, Y. Y., McDonnell, K., Kris, M. G., Shen, R., Sima, C. S., Bach, P. B., Rizvi, N. A., & Riely, G. J. (2010). Pack-years of cigarette smoking as a prognostic factor in patients with stage IIIB/IV nonsmall cell lung cancer. *Cancer*, *116*(3), 670–675. https://doi.org/10.1002/cncr.24813

Jones, J. T., Qian, X., van der Velden, J. L. J., Chia, S. B., McMillan, D. H., Flemer, S., Hoffman, S. M., Lahue, K. G., Schneider, R. W., Nolin, J. D., Anathy, V., van der Vliet, A., Townsend, D. M., Tew, K. D., & Janssen-Heininger, Y. M. W. (2016). Glutathione S-transferase pi modulates NF-κB activation and pro-inflammatory responses in lung epithelial cells. *Redox Biology*, *8*, 375–382. https://doi.org/10.1016/j.redox.2016.03.005

Karin, M. (1999). How NF-kappaB is activated: The role of the IkappaB kinase (IKK) complex. *Oncogene*, *18*(49), 6867–6874. https://doi.org/10.1038/sj.onc.1203219

Khieu, M., & Butler, S. L. (2020). High Grade Squamous Intraepithelial Lesion (HSIL). In *StatPearls*. StatPearls Publishing. http://www.ncbi.nlm.nih.gov/books/NBK430728/

Koetsier, G. (n.d.). *A Practical Guide to Analyzing Nucleic Acid Concentration and Purity with Microvolume Spectrophotometers*. 8.

Liu, T., Zhang, L., Joo, D., & Sun, S.-C. (2017). NF-κB signaling in inflammation. *Signal Transduction and Targeted Therapy*, *2*, 17023. https://doi.org/10.1038/sigtrans.2017.23

Lu, H., Lei, X. & Zhang, Q. Moderate activation of IKK2-NF-kB in unstressed adult mouse liver induces cytoprotective genes and lipogenesis without apparent signs of inflammation or fibrosis. *BMC Gastroenterol* 15, 94 (2015). https://doi.org/10.1186/s12876-015-0325-z

Lung Cancer Statistics | How Common is Lung Cancer. (n.d.). Retrieved January 29, 2020, from https://www.cancer.org/cancer/lung-cancer/about/key-statistics.html

Lousberg, E. L., Fraser, C. K., Tovey, M. G., Diener, K. R., & Hayball, J. D. (2010). Type I Interferons Mediate the Innate Cytokine Response to Recombinant Fowlpox Virus but Not the Induction of Plasmacytoid Dendritic Cell-Dependent Adaptive Immunity. *Journal of Virology*, *84*(13), 6549–6563. https://doi.org/10.1128/JVI.02618-09

MacKinnon, D. P., Fairchild, A. J., & Fritz, M. S. (2007). Mediation Analysis. *Annual Review of Psychology*, *58*, 593. https://doi.org/10.1146/annurev.psych.58.110405.085542

Mascaux, C., Angelova, M., Vasaturo, A., Beane, J., Hijazi, K., Anthoine, G., Buttard, B., Rothe, F., Willard-Gallo, K., Haller, A., Ninane, V., Burny, A., Sculier, J.-P., Spira, A., & Galon, J. (2019). Immune evasion before tumour invasion in early lung squamous carcinogenesis. *Nature*, *571*(7766), 570–575. https://doi.org/10.1038/s41586-019-1330-0

Mazzilli, S. A., Hershberger, P. A., Reid, M. E., Bogner, P. N., Atwood, K., Trump, D. L., & Johnson, C. S. (2015). Vitamin D Repletion Reduces the Progression of Premalignant Squamous Lesions in the NTCU Lung Squamous Cell Carcinoma Mouse Model. *Cancer Prevention Research (Philadelphia, Pa.)*, *8*(10), 895–904. https://doi.org/10.1158/1940-6207.CAPR-14-0403

Merrick, D. T. (2019). Sequencing the Events That Mediate Progression of Premalignant Lung Lesions. *Cancer Research*, *79*(19), 4811–4813. https://doi.org/10.1158/0008-5472.CAN-19-2261

Merrick, D. T., Gao, D., Miller, Y. E., Keith, R. L., Baron, A. E., Feser, W., Kennedy, T. C., Blatchford, P. J., Braudrick, S., Hirsch, F. R., Heasley, L., Bunn, P. A., & Franklin, W. A. (2016). PERSISTENCE OF BRONCHIAL DYSPLASIA IS ASSOCIATED WITH DEVELOPMENT OF INVASIVE SQUAMOUS CELL CARCINOMA. *Cancer Prevention Research (Philadelphia, Pa.)*, *9*(1), 96–104. https://doi.org/10.1158/1940-6207.CAPR-15-0305

Minna, J. D., Roth, J. A., & Gazdar, A. F. (2002). Focus on lung cancer. *Cancer Cell*, *1*(1), 49–52. https://doi.org/10.1016/S1535-6108(02)00027-2

Morabia, A., & Wynder, E. L. (1991). Cigarette smoking and lung cancer cell types. *Cancer*, *68*(9), 2074–2078. https://doi.org/10.1002/1097-0142(19911101)68:9<2074::aid-cncr2820680939>3.0.co;2-x

E. Mortaz, P.A.J. Henricks, A.D. Kraneveld, M.E. Givi, J. Garssen, G. Folkerts, Cigarette smoke induces the release of CXCL-8 from human bronchial epithelial cells via TLRs and induction of the inflammasome, Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease, Volume 1812, Issue 9, 2011, Pages 1104-1110, ISSN 0925-4439, https://doi.org/10.1016/j.bbadis.2011.06.002.

Non-Small Cell Lung Cancer Treatment (PDQ®)–Health Professional Version. (2019, December 13). [PdqCancerInfoSummary]. National Cancer Institute. https://www.cancer.gov/types/lung/hp/non-small-cell-lung-treatment-pdq

Pandya, P. H., Murray, M. E., Pollok, K. E., & Renbarger, J. L. (2016). The Immune System in Cancer Pathogenesis: Potential Therapeutic Approaches. *Journal of Immunology Research*, *2016*. https://doi.org/10.1155/2016/4273943

Pastuszak-Lewandoska, D., Domańska-Senderowska, D., Antczak, A., Kordiak, J., Górski, P., Czarnecka, K. H., Migdalska-Sęk, M., Nawrot, E., Kiszałkiewicz, J. M., & Brzeziańska-Lasota, E. (2018). The Expression Levels of IL-4/IL-13/STAT6 Signaling Pathway Genes and SOCS3 Could Help to Differentiate the Histopathological Subtypes of Non-Small Cell Lung Carcinoma. *Molecular Diagnosis & Therapy*, *22*(5), 621–629. https://doi.org/10.1007/s40291-018-0355-7

Perez-Moreno, P., Brambilla, E., Thomas, R., & Soria, J.-C. (2012). Squamous cell carcinoma of the lung: Molecular subtypes and therapeutic opportunities. *Clinical Cancer Research: An Official Journal of the American Association for Cancer Research*, *18*(9), 2443–2451. https://doi.org/10.1158/1078-0432.CCR-11-2370

Rébé, C., Végran, F., Berger, H., & Ghiringhelli, F. (2013). STAT3 activation: A key factor in tumor immunoescape. *JAK-STAT*, *2*(1), e23010. https://doi.org/10.4161/jkst.23010

Rigden, H. M., Alias, A., Havelock, T., O'Donnell, R., Djukanovic, R., Davies, D. E., & Wilson, S. J. (2016). Squamous Metaplasia Is Increased in the Bronchial Epithelium of Smokers with Chronic Obstructive Pulmonary Disease. *PLoS ONE*, *11*(5). https://doi.org/10.1371/journal.pone.0156009

Singhavi, H., Ahluwalia, J. S., Stepanov, I., Gupta, P. C., Gota, V., Chaturvedi, P., & Khariwala, S. S. (2018). Tobacco carcinogen research to aid understanding of cancer risk and influence policy. *Laryngoscope Investigative Otolaryngology*, *3*(5), 372–376. https://doi.org/10.1002/lio2.204

Squamous Cell Lung Cancer. (2014, July 29). LUNGevity Foundation. https://lungevity.org/for-patients-caregivers/lung-cancer-101/types-of-lung-cancer/squamous-cell-lung-cancer

Suarez, E., & Knollmann-Ritschel, B. E. C. (2017). Squamous Cell Carcinoma of the Lung. *Academic Pathology*, *4*. https://doi.org/10.1177/2374289517705950

Tan, S. C., & Yiap, B. C. (2009). *DNA, RNA, and Protein Extraction: The Past and The Present* [Review Article]. Journal of Biomedicine and Biotechnology. https://doi.org/10.1155/2009/574398

Siegel, R. L., Miller, K. D., & Jemal, A. (2019). Cancer statistics, 2019. *CA: A Cancer Journal for Clinicians*, *69*(1), 7–34. https://doi.org/10.3322/caac.21551

Teixeira, V. H., Pipinikas, C. P., Pennycuick, A., Lee-Six, H., Chandrasekharan, D., Beane, J., Morris, T. J., Karpathakis, A., Feber, A., Breeze, C. E., Ntolios, P., Hynds, R. E., Falzon, M., Capitanio, A., Carroll, B., Durrenberger, P. F., Hardavella, G., Brown, J. M., Lynch, A. G., … Janes, S. M. (2019). Deciphering the genomic, epigenomic, and transcriptomic landscapes of pre-invasive lung cancer lesions. *Nature Medicine*, *25*(3), 517–525. https://doi.org/10.1038/s41591-018-0323-0

The National Lung Screening Trial Research Team. (2011). Reduced Lung-Cancer Mortality with Low-Dose Computed Tomographic Screening. *New England Journal of Medicine*, *365*(5), 395–409. https://doi.org/10.1056/NEJMoa1102873

Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T.-H., Porta-Pardo, E., Gao, G. F., Plaisier, C. L., Eddy, J. A., Ziv, E., Culhane, A. C., Paull, E. O., Sivakumar, I. K. A., Gentles, A. J., Malhotra, R., Farshidfar, F., Colaprico, A., Parker, J. S., … Mariamidze, A. (2018). The Immune Landscape of Cancer. *Immunity*, *48*(4), 812-830.e14. https://doi.org/10.1016/j.immuni.2018.03.023

Westcott, P. M. K., Halliwill, K. D., To, M. D., Rashid, M., Rust, A. G., Keane, T. M., Delrosario, R., Jen, K.-Y., Gurley, K. E., Kemp, C. J., Fredlund, E., Quigley, D. A., Adams, D. J., & Balmain, A. (2015). The mutational landscapes of genetic and chemical models of Kras -driven lung cancer. *Nature*, *517*(7535), 489–492. https://doi.org/10.1038/nature13898

Walser, T., Cui, X., Yanagawa, J., Lee, J. M., Heinrich, E., Lee, G., Sharma, S., & Dubinett, S. M. (2008). Smoking and Lung Cancer. *Proceedings of the American Thoracic Society*, *5*(8), 811–815. https://doi.org/10.1513/pats.200809-100TH

Xia, C., Hu, J., Ketterer, B., & Taylor, J. B. (1996). The organization of the human GSTP1-1 gene promoter and its response to retinoic acid and cellular redox status. *Biochemical Journal*, *313*(Pt 1), 155–161.

Xiong, D., Pan, J., Zhang, Q., Szabo, E., Miller, M. S., Lubet, R. A., You, M., & Wang, Y. (2017). Bronchial airway gene expression signatures in mouse lung squamous cell carcinoma and their modulation by cancer chemopreventive agents. *Oncotarget*, *8*(12), 18885–18900. https://doi.org/10.18632/oncotarget.13806

Yamano, S., Gi, M., Tago, Y., Doi, K., Okada, S., Hirayama, Y., Tachibana, H., Ishii, N., Fujioka, M., Tatsumi, K., & Wanibuchi, H. (2016). Role of deltaNp63posCD44vpos cells in the development of N-nitroso-tris-chloroethylurea-induced peripheral-type mouse lung squamous cell carcinomas. *Cancer Science*, *107*(2), 123–132. https://doi.org/10.1111/cas.12855

Zappa, C., & Mousa, S. A. (2016). Non-small cell lung cancer: Current treatment and future advances. *Translational Lung Cancer Research*, *5*(3), 288–300. https://doi.org/10.21037/tlcr.2016.06.07

CURRICULUM VITAE