Sains Malaysiana 54(3)(2025): 839-868
http://doi.org/10.17576/jsm-2025-5403-18

Potentially Dysregulated Cholesterol, Cellular Interaction, Immune, and Collagen in NTCU-Induced Lung Squamous Cell Carcinoma *in vivo* and LUSC Patients

(Kolesterol Berpotensi Kawal atur, Interaksi Sel, Imun dan Kolagen dalam Karsinoma Sel Skuamosa Paru-paru Aruhan NTCU *in vivo* dan Pesakit LUSC)

MUHAMMAD ASYAARI ZAKARIA[1,4], AMNANI AMINUDDIN[2], NOR FADILAH RAJAB[3], SITI FATHIAH MASRE[1,*] & ENG WEE CHUA[2]

[1]*Centre for Toxicology and Health Risk Studies, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, 50300 Kuala Lumpur, Malaysia*
[2]*Centre for Drug and Herbal Development, Faculty of Pharmacy, Universiti Kebangsaan Malaysia, 50300 Kuala Lumpur, Malaysia*
[3]*Centre for Healthy Ageing and Wellness, Faculty of Health Sciences, Universiti Kebangsaan Malaysia, 50300 Kuala Lumpur, Malaysia*
[4]*Faculty of Pharmacy and Health Sciences, Universiti Kuala Lumpur-Royal College of Medicine Perak, 30450 Ipoh, Perak, Malaysia*

Received: 10 October 2024/Accepted: 26 November 2024

ABSTRACT

Lung squamous cell carcinoma (LUSC) is a deadly cancer, characterized by its complex genetic profiles. Additionally, the molecular mechanisms and etiology underlying LUSC growth are less extensively characterized as compared to adenocarcinoma subtype of lung cancer. Therefore, it is essential to elucidate the molecular mechanisms of LUSC *in vivo* and by using the human database to understand the disease. A LUSC BALB/c mice model was established using N-nitroso-tris-chloroethylurea (NTCU). After termination of mice, the lung tissues were subjected to RNA sequencing, followed by gene set enrichment analysis (GSEA) to identify the enriched pathways. Subsequently, the pathogenic single nucleotide polymorphism (SNP) was determined and enriched using g:Profiler. The transcriptomic profile of human LUSC patients was obtained and analyzed from The International Cancer Genome Consortium (ICGC). The impact of pathogenic simple somatic mutation (SSM) in human LUSC was determined using the Combined Annotation Dependent Depletion (CADD) score, which was also enriched using g:Profiler. Additionally, the enriched pathway of 'Treatment-responsive' was compared with 'Non-responsive' LUSC patients' post-treatment. All pathway analysis was referred to the Reactome database, and an adjusted p-value ≤ 0.05 was considered statistically significant. The top pathway enriched in both mice and human LUSC showed that cholesterol, cellular interaction, immune system, and collagen were significantly affected. Briefly, this study identified important biological pathways that may contribute to LUSC development and hold potential as targets for LUSC therapy in the future.

Keywords: Gene set enrichment analysis (GSEA); lung squamous cell carcinoma (LUSC); RNA sequencing; simple somatic mutation (SSM); single nucleotide polymorphism (SNP)

ABSTRAK

Karsinoma sel skuamus paru-paru (LUSC) adalah kanser yang boleh membawa maut, dicirikan oleh profil genetiknya yang kompleks. Tambahan pula, mekanisme molekul dan etiologi yang mendasari pertumbuhan LUSC adalah kurang dikaji secara mendalam berbanding subjenis adenokarsinoma kanser paru-paru. Oleh itu, adalah penting untuk mengetahui mekanisme molekul LUSC secara *in vivo* dan dengan menggunakan pangkalan data manusia untuk memahami penyakit ini. Model tikus BALB/c LUSC telah dibangunkan menggunakan N-nitroso-tris-kloroetilureum (NTCU). Selepas mencit dikorbankan, tisu paru-paru dianalisis dengan penjujukan RNA, diikuti dengan analisis pengayaan set gen (GSEA) untuk mengenal pasti laluan yang diperkaya. Seterusnya, polimorfisme nukleotida tunggal (SNP) yang patogen telah ditentukan dan diperkaya menggunakan g:Profiler. Profil transkriptomik pesakit LUSC manusia diperoleh dan dianalisis daripada Konsortium Genom Kanser Antarabangsa (ICGC). Kesan mutasi somatik ringkas (SSM) yang patogen dalam LUSC manusia ditentukan menggunakan skor Pengurangan Bergantung Anotasi Gabungan (CADD), yang juga diperkaya menggunakan g:Profiler. Tambahan lagi, laluan yang diperkaya bagi pesakit LUSC 'Responsif-rawatan' dibandingkan dengan pesakit 'Tidak responsif' selepas rawatan. Semua analisis laluan dirujuk kepada pangkalan data Reactome dan nilai p terlaras $\leq$ 0.05 dianggap signifikan secara statistik. Laluan teratas yang diperkaya dalam kedua-dua LUSC model mencit dan manusia

mendedahkan penjejasan signifikan kolesterol, interaksi sel, sistem imun dan kolagen. Ringkasnya, kajian ini mengenal pasti laluan biologi penting yang mungkin menyumbang kepada perkembangan LUSC dan berpotensi dijadikan sebagai sasaran terapi LUSC pada masa hadapan.

Kata kunci: Analisis pengayaan set gen (GSEA); karsinoma sel skuamus paru-paru (LUSC); mutasi somatik ringkas (SSM); penjujukan RNA; polimorfisme nukleotida tunggal (SNP)

INTRODUCTION

Lung cancer is classified into small-cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) histologically and is the leading cause of cancer-related deaths (20%) in both genders in the United States. Each day, approximately 340 people die from lung cancer, which is nearly 2.5 times more than the number of patients who die from colorectal cancer (Siegel, Giaquinto & Jemal 2024). The NSCLC is further divided into three main subtypes: adenocarcinoma, lung squamous cell carcinoma (LUSC), and large cell lung cancer, and accounts for the majority (~85%) of lung cancer cases. Over the past decades, LUSC has gained the attention of researchers globally due to its low five-year overall survival rates of 13% and 2% for stage III and IV patients, respectively (Wang et al. 2020). LUSC patients also tend to develop treatment resistance (Ashrafi et al. 2022; Yi et al. 2024), which may be attributed to the considerably higher oncogenic mutations than other NSCLC subtypes (Gandara et al. 2015; Gu et al. 2023; Yi et al. 2024; Zhao et al. 2016). In addition, LUSC patients are relatively refractory towards the latest therapeutic development for NSCLC, particularly tyrosine kinase inhibitors targeting epidermal growth factor receptors (EGFR) (Liang et al. 2018). This might be due to the low frequency of EGFR amplification (7%) observed in LUSC patients as reported in high-scale genomic analysis (Hammerman et al. 2012). Notably, other molecular aberration-targeted therapies (i.e., ALK, ROS1 & BRAF) also do not benefit LUSC patients due to the difference in genetic profile with other NSCLC subtypes (Korpanty et al. 2014; Tan & Tan 2022). Therefore, a thorough understanding of the molecular profile in LUSC is crucial to understand the disease better and overcoming the plateau advancement of targeted therapies in LUSC.

Advances in nucleic acid sequencing techniques, particularly RNA sequencing, have enabled researchers to profile various cancers and uncover resistance-inducing mutations. The RNA sequencing of lung cancer cells or tissues obtained from pre-clinical models led to the discovery of novel biomarkers and targets for lung cancer precision therapy (Sun et al. 2017; Suzuki et al. 2019; Zilionis et al. 2019). Nevertheless, there is a paucity of pre-clinical lung cancer models for specific LUSC subtypes. Multiple studies have explored methods to induce LUSC *in vivo* reliably and elucidate the pathobiology of LUSC (Henry et al. 1981; Yoshimoto et al. 1980, 1977). One of the most successful techniques is repeated N-nitroso-tris-chloroethylurea (NTCU) treatments in mice (Wang et al. 2004), which have been widely applied by the following researchers and regarded as one of the best pre-clinical LUSC cancer models *in vivo* (Dwyer-Nield et al. 2021; Ghosh et al. 2015; Pan et al. 2018; Riolobos et al. 2019; Surien, Ghazali & Masre 2020; Yamano et al. 2016). The LUSC induction by NTCU is superior to other methods due to the reproducibility and efficacy, besides the similarity of resultant tumors with human LUSC (Surien, Ghazali & Masre 2020; Wang et al. 2009; Zakaria et al. 2021a). Interestingly, LUSC tumors induced by NTCU harbor high genetic profile similarities with the human LUSC, including gene mutations (80%) (Xiong et al. 2018). Hence, it is hypothesized that employing NTCU-induced LUSC in murine models offers a convenient and more accessible approach to unraveling the etiology of human LUSC.

The utilization of NTCU-induced LUSC in mice models has helped researchers gain better insight into the disease. For instance, in a previous study, two LUSC cell lines, UNSCC679 and UN-SCC680, were derived from NTCU-treated mice and analyzed via RNA sequencing. Interestingly, these cells harbor mutated genes relevant to human LUSC, such as G protein-coupled receptor family member apelin receptor (APLNR) and colony-stimulating factor 2 receptor subunit beta (CSF2RB). Moreover, they also found increased expression of genes related to human LUSC from both cell lines, such as *p63*, *Gsto1*, *Aldh3a1*, *Bcl6*, *Atp5g3*, *Dld*, *Odc1*, *Gsta4*, *Ndufb5*, *Ephx1* or *Cox5b* (Valencia et al. 2022). However, this study only focuses on identifying altered genes without comprehensively acknowledging the affected biological pathways. In another study, Pan et al. (2018) performed RNA sequencing on NTCU-induced LUSC, thus, showed the significant enrichment of the PI3K/AKT/NFkB pathways in the disease. This finding further substantiated the role of PI3K in lung cancer development (Cheng et al. 2014). Nevertheless, Pan et al. (2018) performed RNA sequencing on cell samples located along the airway of mice. Briefly, this method was preferred for its robustness, but transcriptomic analysis on excised lung tissues was suggested to be more accurate in depicting potentially overrepresented pathways in human LUSC, as novelly conducted in this study.

Despite the breakthroughs in understanding the molecular mechanisms underlying LUSC, patients still have limited targeted therapy options compared to the adenocarcinoma subtype. This situation may result from a complex and time-consuming drug development process subjected to the approval of the Food and Drug Administration (FDA). Furthermore, the new drug may be ineffective at a clinical trial despite promising results

at the pre-clinical stage. Therefore, researchers should continue investigating the etiology of LUSC, starting from the pre-clinical stage to identify potential drug targets. An animal model that accurately simulates human LUSC, such as NTCU-induced LUSC in mice, remains relevant to accelerate the translation of pre-clinical findings into novel anticancer therapy. This study aimed to provide new insights into the etiology of the NTCU-induced LUSC, focusing on the biological pathways and the impacts of pathogenic mutations by using this model. Furthermore, this study also analyzed the RNA sequencing datasets from the International Cancer Genome Consortium (ICGC) database to identify the impact of pathogenic somatic mutation in LUSC patients and ascertain whether the enriched pathways are associated with their treatment response. In summary, this molecular study provides fundamental knowledge on the underlying molecular events that potentially stimulate LUSC development, thus, helping develop an effective precision therapy for LUSC in the future.

MATERIALS AND METHODS

LUNG SQUAMOUS CELL CARCINOMA (LUSC) DEVELOPMENT *in vivo*

The animal study ethics approval was obtained from Universiti Kebangsaan Malaysia Animal Ethics Committee (FSK/2017/FATHIAH/24-MAY/846-MAY-2017-MAY-2020) and conducted in compliance with the ARRIVE guidelines. All methods were performed following the relevant guidelines and regulations. LUSC in female BALB/c mice model has been developed in our earlier study (Zakaria et al. 2021a). First, 48 female BALBL/c mice (five weeks old), with an average weight of 12-16 g, were purchased from the animal unit, Faculty of Veterinary Medicine, Universiti Putra Malaysia. The mice were divided into two groups: Pre-malignant and Malignant groups. Then, the mice in each group were further allotted into three groups (n=6) according to their respective treatments: 1) Control (receiving 0.9% normal saline), 2) Vehicle (receiving 70% acetone), and 3) Cancer (receiving 0.04 M N-nitroso-tris-chloroethylurea (NTCU)). Before treatment, all mice were acclimatized at room temperature for two weeks with *ad libitum* access to regular tap water and mice pellet. The 0.04 NTCU is a nitrosamine-derivative carcinogen for inducing LUSC in mice, which was dissolved in 70% acetone to obtain the respective concentration. 25 µL of 0.9% normal saline, 70% acetone, and 0.04M NTCU to each respective group was administered at the dorsal area, between the shoulder blade bones of mice shaved skin (2 cm^2 area) via skin painting. The painting was performed by spreading the chemical using pipette tips and performed twice per week with an interval of 3.5 days for 30 weeks. At the end of 30 weeks after the initial treatment, the animals were terminated using an overdose of ketamine-xylazine cocktail (KTX) and cervical dislocation. The KTX dose used to euthanize the mouse in this study was 0.2 mL/20 mg of mice. KTX was administered via intraperitoneal injection. Verification of death was upon the absence of reflex following the pinching of the mouse paw. The histology analysis performed in our earlier study (Zakaria et al. 2021a), combining H&E and IHC staining, confirmed the successful induction of pre-malignant and malignant LUSC in our research. For the RNA sequencing performed in this study, we analyzed the tissues from the malignant vehicle and cancer groups. We analyzed the vehicle group only since there was no significant difference in histological characteristics, histopathology scores, and epithelium thickness between the control and vehicle groups (Zakaria et al. 2021a). We also analyzed malignant groups only to gain insight into the overall dysregulated pathway responsible for LUSC carcinogenesis.

RNA EXTRACTION AND QUALITY CHECKS

At the end of the animal study, approximately 20 mg of lung tissue was excised from each mouse, immersed in lysis buffer, and homogenized using ULTRA-TURRAX® T25 Basic (IKA, Germany). Subsequently, the RNA extraction was performed using the innuPREP RNA Mini Kit (Analytik Jena, Germany) according to the manufacturer's guidelines. As a precaution, contaminating genomic DNA in the samples was removed using the RapidOut DNA Removal Kit (Thermo Scientifc Inc., USA) as instructed by the manufacturer. Finally, the concentration and purity of the RNA samples were measured via the OPTIZEN™ NanoQ Microvolume Spectrophotometer (Mecasys Co. Ltd, Korea), and the RNA integrity was evaluated using 1% agarose gel electrophoresis (Supplementary Figure S1). All samples that passed the quality checks outlined by the company (absorbance ratios = 1.8 - 2.1 at 260 nm and 280 nm wavelength, exhibited discrete 28S and 18S bands) were subjected to transcriptomic analysis.

RNA SEQUENCING

The integrity (RIN = RNA integrity number) of RNA samples was evaluated using Agilent 2100 Bioanalyzer (Agilent Technologies Inc., USA) by Apical Scientific Sdn. Bhd., Malaysia. As all the samples recorded good RNA integrity (RIN > 6.8) (Supplementary Figure S2), the samples were then proceeded to RNA sequencing on a HiSeq150 platform (Illumina, USA) by Novogene Corporation Inc., China.

BIOINFORMATIC ANALYSIS

Once the sequencing was completed, the raw reads were obtained and filtered to remove unwanted reads: adaptors, undeterminable bases (N > 10%), and reads with > 50% bases (quality scores ≤ 5). Subsequently, the clean reads were aligned to the *Mus musculus* genome

(GRCm38/mm10) using STAR version 2.5 (Dobin et al. 2013). Gene-level read counts were then computed using HTSeq version 0.6.1 (Anders, Pyl & Huber 2015), followed by the normalization across all samples to eliminate the unwanted technical variation between samples using the DESEq2 version 2_1.6.3 (Love, Huber & Anders 2014). The principal component analysis (PCA) plot analysis showed that the Vehicle and Cancer groups have distant gene expression patterns (Supplementary Figure S3). After that, Gene Set Enrichment Analysis (GSEA) and the functional impact of the DNA mutations were analyzed. Raw and processed sequenced data were submitted to the GEO database (accession ID: GSE240837).

GENE SET ENRICHMENT ANALYSIS (GSEA) OF MOUSE LUSC

The normalized read counts were analyzed using GSEA version 4.2.3 (http://www.broadinstitute.org/gsea/downloads.jsp) to pinpoint differentially expressed genes and pathways (Mootha et al. 2003; Subramanian et al. 2005). Reactome 'c2.cp.reactome.v2022.1.Hs.symbols.gmt [Curated]' with the 'Mouse_ENSEMBL_Gene_ID_Human_Orthologs_MSigDB.v2022.1.Hs.chip' chip platform was the gene sets database utilized in this study (Liberzon et al. 2011). The statistical significant pathways were regarded if FDR < 25%, normalized enrichment score (NES) > 1, and nominal $p < 0.05$. At the end of the analysis, NES, which reflects the extent of a gene set overrepresentation, and the enriched pathways were ranked in descending order. Finally, the list of genes affected by the top ten pathways was identified using 'leading edge analysis' in GSEA to identify specific genes that strongly contribute to the detected pathways.

SINGLE NUCLEOTIDE POLYMORPHISM (SNP) MUTATIONAL ANALYSIS OF MOUSE LUSC

The impact of single nucleotide polymorphism (SNP) mutations was analyzed in this study, while the effect of insertion-deletion (Indels) mutations was excluded. Most variants (90%) showed in this study were SNP mutations, and the pathogenicity specificity was enhanced by using two prediction algorithms. The variant calling was performed using Genome Analysis Toolkit (GATK) version 3.0 (McKenna et al. 2010), followed by the prediction of functional impacts of the resultant list of mutations using a combination of Sorting Intolerant From Tolerant (SIFT) and Protein Variation Effect Analyzer (PROVEAN) algorithms via a web-based server at http://provean.jcvi.org/index.php (Choi et al. 2012). Pathogenic mutations are defined by SIFT scores of < 0.05 and PROVEAN scores of < - 2.5. Subsequently, the list of affected genes was extracted from the genomic coordinates of the DNA variants. BioVenn (https://www.biovenn.nl/) (Hulsen, de Vlieg & Alkema 2008), was later utilized to produce a list of mutated genes unique to the cancer group, as shown in the Supplementary Table S1. The data were then analyzed using g:Profiler (https://biit.cs.ut.ee/gprofiler/gost) (Raudvere et al. 2019) to identify overrepresented pathways.

ANALYSIS OF DATASETS FROM THE CANCER GENOME ATLAS PROGRAM (TCGA)

The TCGA datasets from the ICGC database for LUSC patients were analyzed to corroborate the RNA sequencing findings. The TCGA-LUSC cohort comprised 428 patients (project code, LUSC-US) with documented outcomes of their last follow-up assessments (complete remission, n = 229; progression, n = 15; deceased, n = 137; or unknown, n = 47). The patients were then classified into two major groups to evaluate potential pathways associated with treatment responsiveness: 'Treatment-responsive' = LUSC patients with complete remission and 'Non-responsive' = patients who suffered disease progression or were deceased (Supplementary Table S2). Patients with unknown status were omitted. Once the classification was completed, the GSEA was used to identify enriched pathways in LUSC patients with good therapeutic responses. The Reactome pathway database, 'c2.cp.reactome.v2022.1.Hs.symbols.gmt [Curated]' with the 'Human_NCBI_Gene_ID_MsigDB.v2022.1.Hs.chip' chip platform in GSEA was utilized for this purpose (Liberzon et al. 2011).

Simple somatic mutations (SSMs) were also evaluated in this study. In-house R scripts (Supplementary Data S1) were used to convert datasets (tab-separated values) downloaded from ICGC into VEP-compatible formats, remove duplicates, add functional annotations, and compute aggregate scores to reflect gene-level mutational burdens. First, the SSMs were filtered for principal transcripts documented in the Annotation of Principal and Alternative Splice Isoforms (APPRIS) database (http://appris.bioinfo.cnio.es) (Rodriguez et al. 2013). Second, the SSMs with Combined Annotation Dependent Depletion (CADD) scores of > 20, a cut-off point commonly used to define pathogenic DNA variants, were selected (Kircher et al. 2014) and accumulated for all patients to obtain gene-level scores. Finally, the top 25% of mutated genes were ranked in descending order of the aggregate CADD scores and analyzed using g:Profiler using the 'Ordered query' option.

STATISTICAL ANALYSIS

Differential expression analysis between the Vehicle and Cancer groups was performed using the DESeq2 version 2_1.6.3. The resulting p-values were adjusted using Benjamini and Hochberg's approach for controlling the False Discovery Rate (FDR), and genes with an adjusted p-value or $p$adj value ≤ 0.05 were assigned as differentially expressed. For the enrichment analysis using g:Profiler, the enriched terms in Gene Ontology (GO) or Reactome database were considered significant if the $p$adj ≤ 0.05.

RESULTS

THE GSEA OF LUSC VERSUS NORMAL LUNG TISSUES OBTAINED FROM MICE

Based on the enrichment analysis using annotations curated in the Reactome database, the top five pathways enriched in the mice LUSC using GSEA were identified: cholesterol biosynthesis, keratinization, activation of gene expression by Sterol regulatory element-binding transcription factor (SREBF), Sterol Regulatory Element-Binding Protein (SREBP), formation of the cornified envelope, and neutrophil degranulation. Cholesterol biosynthesis was the most critical (enriched) contributor to LUSC (NES = 2.6327), followed by keratinization (NES = 2.5642) and activation of gene expression by SREBF (SREBP) (NES = 2.5573) (Figure 1).

LEADING EDGE ANALYSIS OF THE ENRICHED PATHWAYS IN THE LUSC MOUSE MODEL

The top 20 genes affected from the top 10 most enriched pathways were identified from GSEA via the leading-edge analysis: *KRT1*, *PKP1*, *HMGCS1*, *CYP51A1*, *KRT4*, *CFP, FASN, PGAM1, PTAFR, KRT18, KRT6C, PKP2, TCHH, FURIN, IGHV3-23, LGALS3, ANPEP, SLC11A1, SCAMP1,* and *SLPI* (Figure 2).

PATHWAYS AFFECTED BY SINGLE-NUCLEOTIDE POLYMORPHISMS (SNP) IN MOUSE LUSC

The pathways affected by pathogenic SNP in mouse LUSC were assessed using g:Profiler (Figure 3). Notably, no pathways from the Reactome database were enriched. However, gene ontology analysis yielded several significantly enriched terms. Three biological processes, three molecular functions, and four cellular components were significantly enriched due to pathogenic SNP in LUSC (Figure 3). From these gene ontology results, two specific terms were identified, namely homophilic cell adhesion via plasma membrane adhesion molecules and pyrroline-5-carboxylate reductase activity.

THE GSEA OF TREATMENT-RESPONSIVE VERSUS NON-RESPONSIVE LUSC PATIENTS

In comparing the output from LUSC patients who responded well to treatments against those who did not, the top five pathways associated with treatment responsiveness were identified: nuclear signaling by ERBB4, signaling by ERBB4, inactivation of CSF3 (G-CSF) signaling, TNF receptor-associated factor 6 (TRAF6)-mediated induction of TGFβ-activated kinase 1 (TAK1) complex within Toll-like receptor 4 (TLR4) complex, and interleukin 10 (IL-10) signaling. Nuclear signaling by ERBB4 was the most enriched pathway in patients with treatment-responsive LUSC (NES = 1.8345), followed by signaling by ERBB4 (NES = 1.7905), and inactivation of CSF3 (G-CSF) signaling (NES = 1.658) (Figure 4).

LEADING EDGE ANALYSIS OF ENRICHED PATHWAYS IN RESPONSIVE HUMAN LUSC TOWARD TREATMENTS

Similar to the leading-edge analysis performed from enriched pathways in mice LUSC, genes affected from the top 10 most enriched pathways in responsive human LUSC toward treatments were also identified. The top 20 genes are *STAT5A, UBC, SOCS3, TAB2, SPARC, CSN2, PGR, PSEN2, APOE, CXCL12. HCK, HBEGF, WWOX, CD86, CSF1, TNF, MAP3K7, TICAM2, CISH,* and *DEFB115* (Figure 5).

FUNCTIONAL IMPACT OF SIMPLE SOMATIC MUTATIONS (SSM) IN HUMAN LUSC

The pathogenic SSM analysis on the human LUSC yielded 41 overrepresented pathways. The top five pathways affected are collagen chain trimerization, MET activates PTK2 signaling, ECM proteoglycans, L1-ankyrins interaction, and assembly of collagen fibrils and other multimeric structures (Figure 6).

The overrepresented pathways and mutational analysis results obtained from mice and human data can be categorized into four major themes: cholesterol, cellular interaction, immune system, and collagen dysregulation in LUSC. Cholesterol dysregulation may be attributed to cholesterol biosynthesis and activation of gene expression by SREBF (SREBP). Meanwhile, cellular interaction may be dysregulated through homophilic cell adhesion via plasma membrane adhesion molecule, MET activates PTK2 signaling, and interaction between L1 and ankyrin. Immune dysregulation may involve neutrophil degranulation, CSF3 (G-CSF) signaling, TRAF6-mediated induction of TAK1 complex within TLR4 complex, and IL-10 signaling pathways. Finally, ECM dysregulation may involve pyrroline-5-carboxylate reductase (PYCR) activity, collagen chain trimerization, and assembly of collagen fibrils and other multimeric structures.

DISCUSSION

LUSC is the second most common subtype of lung cancer after adenocarcinoma (Alipour et al. 2024), associated with a poor therapeutic response due to the distinct genetic profile from other lung cancer subtypes (Anusewicz, Orzechowska & Bednarek 2020; Zhang et al. 2019). Thus, it is crucial to elucidate the underlying molecular aberrations and overrepresented pathways that potentially contribute to LUSC development. The present study investigated lung tumors' transcriptomic and mutational profiles excised from NTCU-treated mice to understand LUSC pathobiology better. In addition, the current findings were cross-validated with RNA sequencing datasets of LUSC patients from the ICGC database.

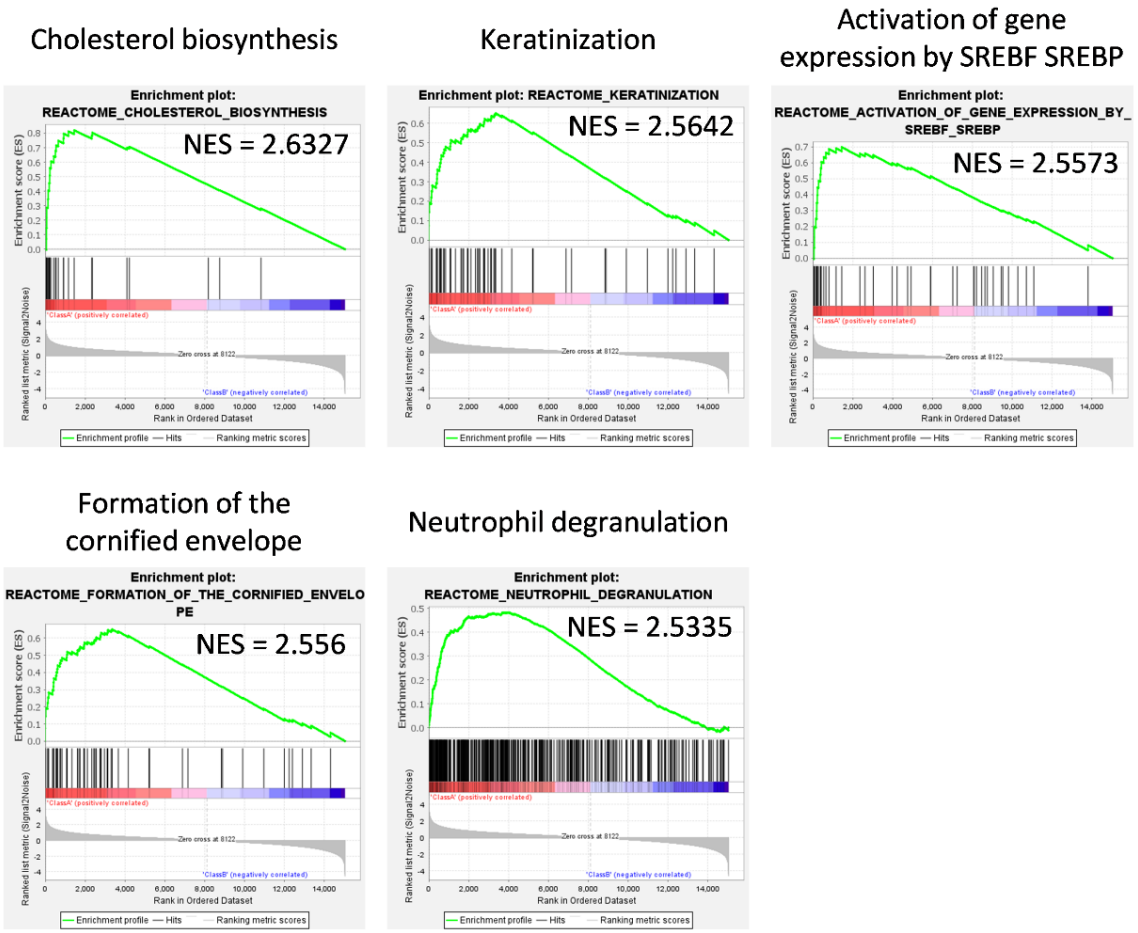


FIGURE 1. The enrichment plots illustrate the top five pathways enriched in the mice LUSC. The pathways were analyzed using GSEA in descending order: cholesterol biosynthesis, keratinization, activation of gene expression by SREBF (SREBP), formation of the cornified envelope, and neutrophil degranulation

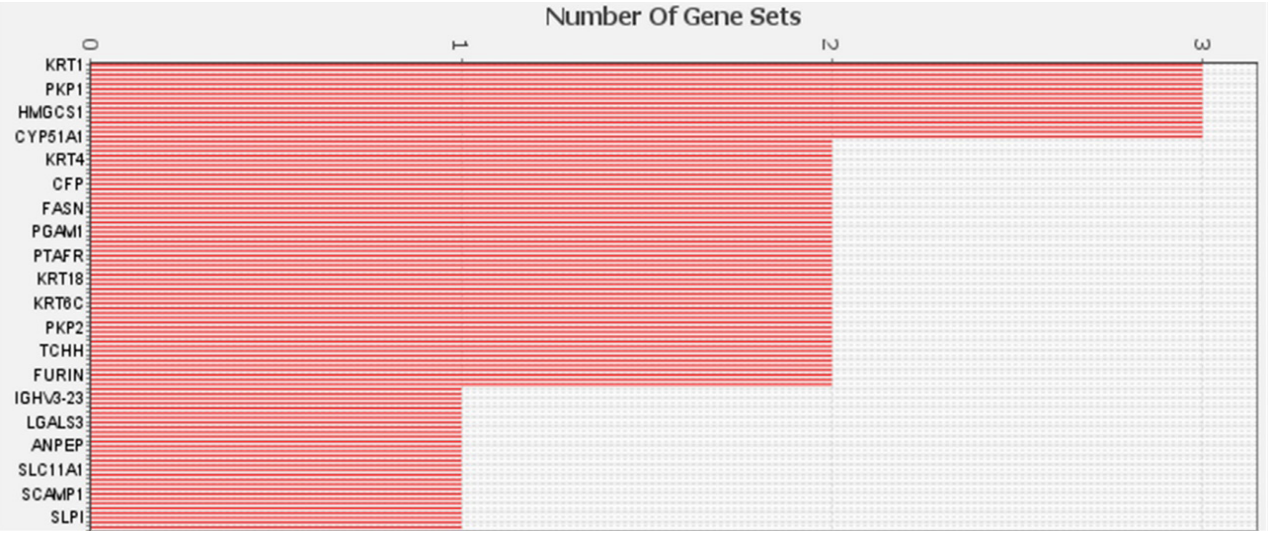


FIGURE 2. The top 20 genes affected in the top 10 most enriched pathways from the mice LUSC. The genes were identified using leading-edge analysis in GSEA

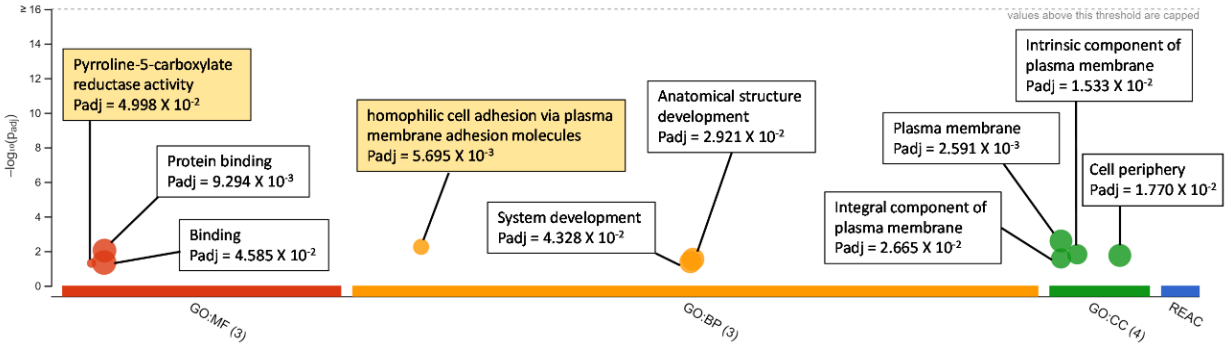


FIGURE 3. The significantly enriched gene ontology terms impacted by the pathogenic SNP mutation in the mice LUSC. MF = molecular functions, BP = biological processes, and CC = cellular components. The pathways were labeled accordingly with their respective padj value, while the shaded text boxes indicated the two specific terms

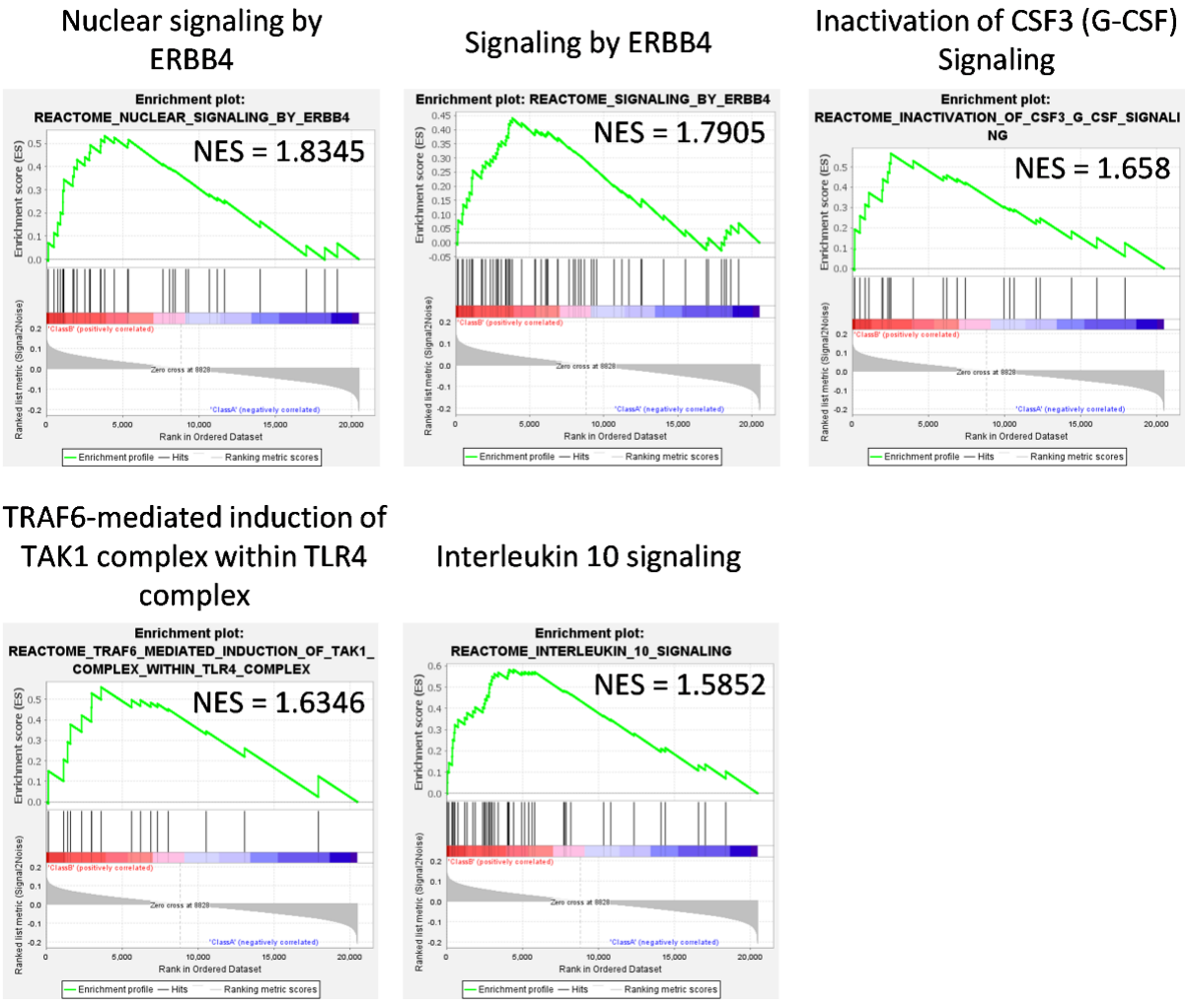


FIGURE 4. The enrichment plots illustrate the top five pathways enriched in treatment-responsive LUSC patients, analyzed in GSEA. The top five pathways in descending order are nuclear signaling by ERBB4, signaling by ERBB4, inactivation of CSF3 (G-CSF) signaling, TRAF6-mediated induction of TAK1 complex within TLR4 complex, and IL-10 signaling

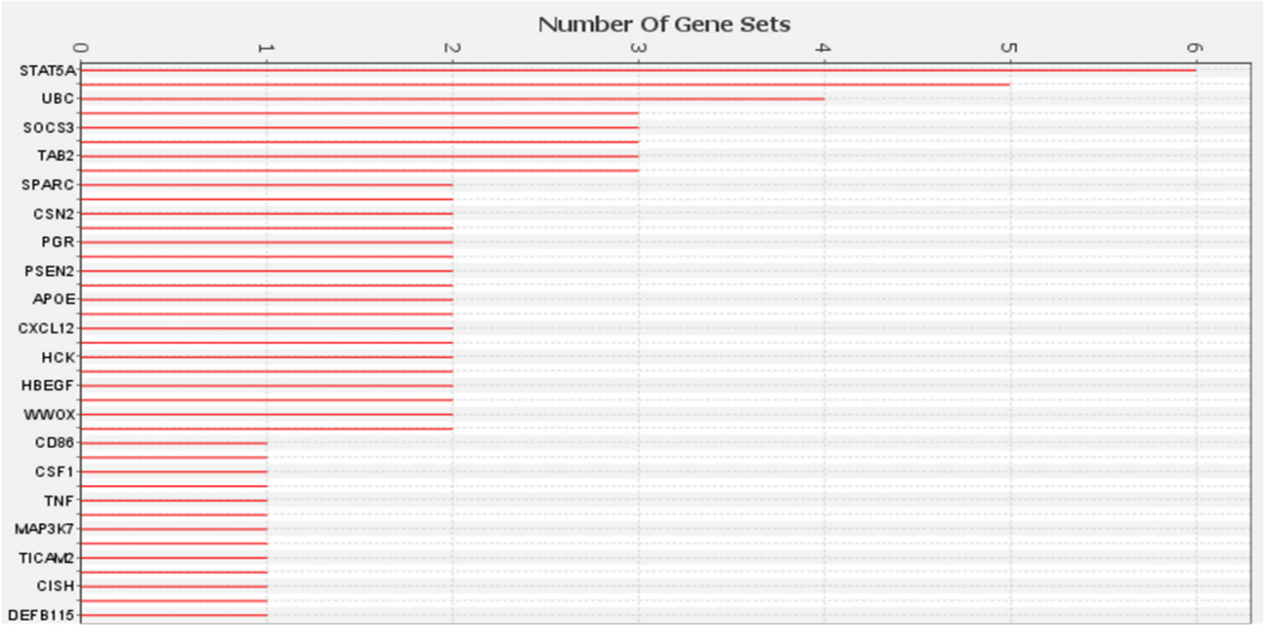


FIGURE 5. The top 20 genes affected in the top 10 most enriched pathways from treatment-responsive LUSC patients. The genes were identified using leading-edge analysis in GSEA

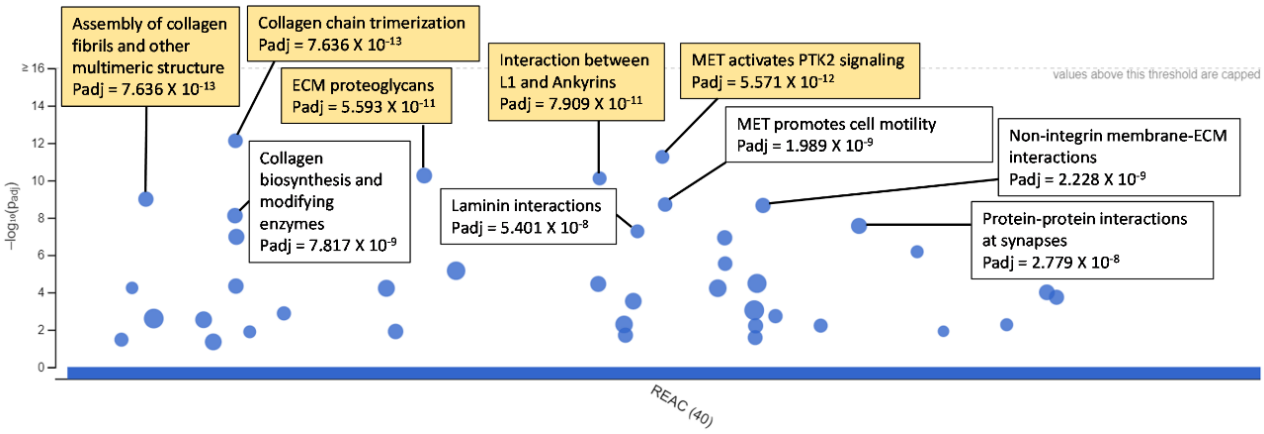


FIGURE 6. The significantly enriched pathways impacted by the pathogenic SSM mutation in human LUSC, curated from the Reactome database. The top 10 pathways were labeled accordingly with their respective padj value, while the shaded text boxes indicated the top five pathways

Based on the transcriptomic analysis of mice LUSC, cholesterol was identified as one of the main dysregulated terms. Our finding aligns with previous studies on increased cholesterol biosynthesis in the lung adenocarcinoma subtype (Hartmann et al. 2023; Hoppstädter et al. 2021). Cholesterol plays a vital role in maintaining healthy cellular membrane composition, synthesis of hormones, and upregulation of transducing signaling pathways, which has been reported in cancer (Mok & Lee 2020). In particular, increased cholesterol biosynthesis is also necessary to meet the high demand for constructing new cell membranes in actively dividing cancer cells (Coradini, Ambrogi & Infante 2023). Notably, cholesterol may help cancer growth by assisting cancer cells in apoptosis evasion, migration, and metastasis (Gu et al. 2019; Jiang et al. 2019; Kuzu et al. 2017). Cholesterol also triggers the activation of oncogenic signaling pathways, such as Hedgehog and mTORC1 (Castellano et al. 2017; Huang

et al. 2016). Interestingly, the enriched activation of gene expression by SREBF (SREBP) is remarkably associated with cholesterol biosynthesis. SREBP acts as a master transcription factor that can induce lipogenic enzyme activity, required in mediating cholesterol biosynthesis (Foretz et al. 1999; Shimano et al. 1999). SREBP has been shown to promote human glioblastoma multiforme (GBM) cancer cell viability through cholesterol accumulation, and its inhibition has been proven to impair cancer cell survival under hypoxia conditions (Lewis et al. 2015). Thus, our finding suggests that LUSC may follow the common motive of other types of cancer to survive by upregulating cholesterol biosynthesis and SREBP.

The dynamics of cellular interaction have long been recognized to play a pivotal role in driving metastasis in cancer (Janiszewska, Primi & Izard 2020). In our current study, cellular interaction was possibly dysregulated in LUSC. Notably, the homophilic cell adhesion via plasma membrane adhesion molecules was significantly affected. This biological process involves the attachment of a plasma membrane adhesion molecule in one cell to an identical molecule in an adjacent cell. In particular, the enrichment dataset highlighted cadherin and nephrin as the affected proteins required for cell-cell communication and sustaining tissue integrity (Khoshnoodi et al. 2003; Kourtidis et al. 2017). Mounting evidence corroborates the impact of the gain or loss of cadherin binding on cancer growth. For instance, E-cadherin downregulation and N-cadherin upregulation can induce epithelial-mesenchymal transition (EMT) (Yu et al. 2015), hence enabling cancer cells to migrate and metastasize (Tsai & Yang 2013). Furthermore, the MET activates PTK2 signaling, which could also be associated with regulating cellular interaction. The MET is a tyrosine kinase receptor vital for cell proliferation, differentiation, and migration (Taghehchian et al. 2021). Meanwhile, protein tyrosine kinase 2 (PTK2) is one of the downstream targets for MET. Its activation was linked to cell adhesion regulation and cytoskeleton rearrangement, impacting cancer cell behavior (Zakaria et al. 2021b). Similarly, the interaction between L1 and ankyrin could also be involved in the cell adhesion regulatory process. Neural cell adhesion molecule L1 (L1) is a transmembrane glycoprotein expressed on the cell surface. This glycoprotein can bind with a cytoskeleton family protein known as ankyrin, which links the plasma membrane with the internal cytoskeleton (Cau et al. 2022). The dysregulation of L1-ankyrin interaction can enhance cell attachment and detachment from the neighboring cells or ECM substrate, consequently facilitating cell migration, an initial step for metastasis (Dou et al. 2018). In summary, our study shows potential cellular interaction dysregulation, exemplified by homophilic cell adhesion, MET-PTK2, and L1-ankyrin interactions.

The immune system is an essential component of the human body that ironically plays a diverse role in cancer development (Gonzalez, Hagerling & Werb 2018). Besides cancer cells and other stromal cells, the tumor microenvironment is also home to infiltrating immune cells such as neutrophils, macrophages, monocytes, mast cells, natural killer (NK) T-cells, B cells, a cluster of differentiation (CD) 4 T-cells, and CD 8 T-cells (Bhargav et al. 2023). Neutrophils, for instance, can be recruited by cancer cells to promote multiple hallmarks of cancer, such as immunosuppression and normal tissue remodeling into tumors (Mollinedo 2019). Moreover, neutrophils can induce angiogenesis by releasing matrix metalloproteinase-9 (MMP-9). Neutrophils can promote the dissemination of cancer cells to the surrounding tissues by releasing heparanase and neutrophil elastase via degranulation (Ardi, Deryugina & Quigley 2020; Mayfosh, Baschuk & Hulett 2019). This process occurs when neutrophil cytoplasmic granules fuse with the cell membrane, thus resulting in soluble granule protein exocytosis. Therefore, the process of neutrophil degranulation enriched in the present LUSC model may be required for cancer progression, as proven by earlier studies in different types of cancer (Emmons et al. 2021; Strell et al. 2010).

Increasing evidence has underscored the role of collagen in promoting lung cancer carcinogenesis. Based on the mutational analysis in LUSC tissues *in vivo*, this study showed that pyrroline-5-carboxylate reductase (PYCR), an enzyme that converts Δ1-pyrroline-5-carboxylate (P5C) into proline necessary for collagen biosynthesis (Christensen et al. 2017), was impacted in the present study. This enzyme is highly expressed in different cancer types and can influence cancer cells' behavior, such as proliferation, migration, and therapeutic response (Wang & Liu 2019; Wang et al. 2019; Weijin et al. 2019; Zeng et al. 2017). Earlier studies demonstrated that PYCR knockdown inhibited NSCLC cell proliferation and cell cycle in lung cancer (Cai et al. 2018), besides hampering cellular EMT (Sang, Zhang & Shan 2019). Likewise, the sensitivity of lung adenocarcinoma cells to cisplatin was enhanced upon PYCR knockdown (She et al. 2019), suggesting the critical role of this enzyme in nurturing resistance to treatment and promoting cancer growth. As an impact of PYCR dysregulation, collagen was also found to be highly expressed in LUSC from this study. Collagen is a central component of the ECM in normal tissue, responsible for sustaining the interstitial structure that affects physical tissue characteristics (tissue rigidity) (Bordeleau et al. 2017; Xu et al. 2019; Zakaria et al. 2022). Collagen support in carcinogenesis is executed in multiple ways. For instance, this protein promotes immunosurveillance by preventing the entry of CD8+ cytotoxic T-cells (primary lymphocytes tasked to eliminate cancer cells) into cancer tissue (Voiles et al. 2014). Moreover, collagen promotes cancer cell proliferation, invasion, and metastasis by interacting with other ECM molecules (Liu et al. 2018; Natarajan et al. 2019; Xu et al. 2019). Consistent with our identification of enriched collagen-associated terms, earlier studies have also reported elevated collagen levels in blood samples of

NSCLC patients and 3D-cell cultures containing NSCLC cell lines (Fang et al. 2019; Tamiya et al. 2013; Voiles et al. 2014). Briefly, the crucial role of collagen in LUSC might be reinforced by enriched PYCR and collagen in our study. However, this study only screened the mutational impacts mentioned, and further experimental validation is required to confirm the effects of PYCR and collagen dysregulation in LUSC progression.

The data analysis of treatment-responsive LUSC patients yielded several enriched signaling pathways, such as ERBB4, CSF3 (G-CSF), TRAF6-mediated induction of the TAK1 complex within the TLR4 complex, and IL-10 signaling. These pathways were consistent with the major themes enriched in mice and human LUSC of our study, thus, potentially associated with LUSC carcinogenesis. Although some pathways may serve as good LUSC prognoses (CSF-3 inactivation and IL-10 signaling), the underlying mechanism of the top five pathways is linked with four major themes of pathways enriched in LUSC, which offers exciting areas for future research.

The ERBB4 protein, a member of the EGFR family, is crucial in cell proliferation and survival (Gong et al. 2020; Williams et al. 2015). Mechanistically, this receptor protein can dimerize upon ligand stimulation, activating various signaling pathways, such as Ras/Raf/MAPK and PI3K/AKT, that promote cell proliferation (Guenzi et al. 2021). Notably, overexpression of ERBB4 mRNA in lung adenocarcinoma patients has been linked to distant metastases, advanced TNM stages, and poor overall survival, indicating its prognostic significance (Masroor et al. 2016). The gene that codes this protein was positioned among the top 30 most frequently altered genes in lung cancer: (26^{th} place in adenocarcinoma), (22^{nd} place in LUSC), and (6^{th} place in small cell lung cancer) (Fang et al. 2015). According to Ding et al. (2008), The ERBB4, along with other tyrosine kinase genes, was suggested as a proto-oncogene in lung cancer, but can also function as a tumor suppressor in its homodimer form (Lucas et al. 2022). An earlier study also reported a better prognosis of advanced NSCLC patients with ERBB4 mutation when treated with immune checkpoint inhibitors than those with ERBB4 wild type (Hu et al. 2021). This may occur due to increased sensitivity to chemotherapeutic agents such as Lapatinib in cancer that harbor mutant ERBB4, as discussed by Lau et al. (2014). Collectively, these findings underscore the importance of understanding ERBB4 expression levels in lung cancer patients for prognostic evaluation and potential therapeutic interventions.

The inactivation of Colony-stimulating factor-3 (CSF-3) was also enriched in treatment-responsive LUSC patients. CSF-3 could promote tumorigenicity through angiogenesis, chemotherapeutic resistance, and apoptosis inhibition (Kawano et al. 2015; Okazaki et al. 2006). This glycoprotein was highly expressed in tumors compared to normal tissues of the same organ (Liu et al. 2020; Morris et al. 2014). Therefore, CSF-3 inactivation observed in treatment-responsive patients is rational as it may improve the prognosis of LUSC patients, which aligned with a previous study that reported an increased T-cell infiltration and decreased tumor growth of colorectal cancer following CSF-3 inhibition (Morris et al. 2015). The involvement of CSF-3 in regulating immune reaction also agrees with our findings that suggest immune dysregulation as one of the significant themes enriched in LUSC. The CSF-3 was reported to be among the critical CSFs required for neutrophil production (Hamilton & Achuthan 2013). Moreover, CSF-3 plays a vital role in the tumor microenvironment by attracting immune cells, especially neutrophils to the tumor site. These cells then undergo reprogramming and contribute to angiogenesis (Missiaen, Mazzone & Bergers 2018), which can aid in cancer cell dissemination and metastasis (Kowanetz et al. 2010).

The TRAF6-mediated induction of the TAK1 complex within the TLR4 complex and IL-10 was also associated with immune modulation in cancer. The TRAF6 can stimulate the TAK1 activation (Li et al. 2020), leading to the activation of multiple pathways, including the TLR4 signaling pathway (Kim et al. 2022). The TLR4 is crucial for innate immunity by regulating the NF-κB transcription factor (Li et al. 2020), which mediates inflammatory reactions by neutrophil recruitment in innate immune response (Verstrepen et al. 2008). Interestingly, TLR4/NF-κB inhibition has been shown to reduce inflammation-induced oxidative stress in acute lung injury (Zhang et al. 2019), suggesting the importance of the TRAF-6/TAK1/TLR pathway in lung pathogenesis. IL-10 is another pathway enriched in treatment-responsive LUSC, which warrants attention due to its significant role in shaping the tumor microenvironment. IL-10 is an anti-inflammatory cytokine that can promote cancer growth by enabling immunosurveillance (Dennis et al. 2013). Notably, IL-10 can be secreted by neutrophils, which were enriched in our study. IL-10 is also among the earliest cytokines produced by the myeloid cell in acute lung disease (González et al. 2021), suggesting its essential role in lung pathology, including LUSC. In brief, the TRAF-6 mediated activation of TAK1 within the TLR4 complex and IL-10 were suggested to be essential for immune modulation in LUSC.

Several limitations have been identified in this study. First, low-frequency transcripts were likely undetectable at a sequencing depth of ~20 million reads per sample. CADD annotations were unavailable to assess the functional relevance of SNPs in mouse LUSC, which can lead to the potential missing of several critical effects from Indels. Moreover, the complete TCGA dataset could not be evaluated, possibly leading to incomprehensive findings on the functional biological themes. For future studies, it is suggested that *in vitro* and *in vivo* assays be performed to confirm the involvement of these four themes in LUSC carcinogenesis. Using transgenic mice with modifications to several genes identified in this study would help clarify their impact on LUSC. It is also recommended that a lung

adenocarcinoma group be included in future studies to elucidate further the differences in the four major themes identified between LUSC and lung adenocarcinoma. Collectively, our suggestion for future studies may help the researcher better understand potential biomarkers, target genes, and signaling pathways associated with LUSC. Moreover, researchers can identify the unique themes or independent factors that may be possessed by specific subtypes of lung cancer, eventually leading to a big leap toward personalized medicine.

CONCLUSION

The study findings provide insights into the mechanisms underlying LUSC and potential therapeutic targets and predictive biomarkers in response to treatment from LUSC patients. Resultantly, cholesterol-, cellular interaction-, immune system-, and collagen-associated terms were enriched in mouse and human LUSC. Herein, this study suggested the potential contribution of cholesterol, SREBP, homophilic cell adhesion, PYCR, and ECM proteins, such as collagen, in mediating LUSC. Furthermore, the pathways enriched in the treatment-responsive LUSC patients could be associated with the pathobiology of LUSC. The potential target genes and pathways discussed are worth exploring as promising avenues for therapeutic intervention of LUSC. Lastly, this study proposed NTCU-induced LUSC as a feasible model to understand LUSC carcinogenesis at a molecular level due to the similarities in several pathways or themes with human LUSC.

ACKNOWLEDGMENTS

This work was supported by the Ministry of Higher Education under the Fundamental Research Grant Scheme (FRGS/1/2018/SKKKK06/UKM/03/1) and Dana Impak Perdana (DIP-2016-017). We would like to thank the Science Officer; Miss Mazlin Aman. Associate Science Officer; Madam Azlina Izni Mohd Nor and laboratory technologist; Madam Nurulfatihah Samuri from the Pathology Radical laboratory, Faculty of Health Sciences, Universiti Kebangsaan Malaysia. Additionally, the would also like to thank the science officer; Madam Hazni Falina Mohamad and laboratory technologists; Madam Azlina Razimahwati Azizi from the Faculty of Pharmacy, Universiti Kebangsaan Malaysia, for their assistance throughout the research.

REFERENCES

Alipour, M., Moghanibashi, M., Naeimi, S. & Mohamadynejad, P. 2024. Integrative bioinformatics analysis reveals ECM and nicotine-related genes in both LUAD and LUSC, but different lung fibrosis-related genes are involved in LUAD and LUSC. *Nucleosides, Nucleotides & Nucleic Acids* 43(9): 915-934.

Anders, S., Pyl, P.T. & Huber, W. 2015. HTSeq - a Python framework to work with high-throughput sequencing data. *Bioinformatics* 31(2): 166-169.

Anusewicz, D., Orzechowska, M. & Bednarek, A.K. 2020. Lung squamous cell carcinoma and lung adenocarcinoma differential gene expression regulation through pathways of Notch, Hedgehog, Wnt, and ErbB signalling. *Scientific Reports* 10(1): 21128.

Ardi, V.C., Deryugina, E.I. & Quigley, J.P. 2020. Neutrophil elastase facilitates cancer cell motility via induction of Akt signaling pathway. *The FASEB Journal* 34(S1): 1-1.

Ashrafi, A., Akter, Z., Modareszadeh, P., Modareszadeh, P., Berisha, E., Alemi, P.S., Chacon Castro, M.D.C., Deese, A.R. & Zhang, L. 2022. Current landscape of therapeutic resistance in lung cancer and promising strategies to overcome resistance. *Cancers* 14(19): 4562.

Bhargav, A., Bhalla, N., Manoharan, S., Singh, G., Yadav, S.K. & Singh, A.K. 2023. Role of various immune cells in the tumor microenvironment. *Diseases & Research* 3(1): 30-40.

Bordeleau, F., Mason, B.N., Lollis, E.M., Mazzola, M., Zanotelli, M.R., Somasegar, S., Califano, J.P., Montague, C., LaValley, D.J., Huynh, J., Mencia-Trinchant, N., Negrón Abril, Y.L., Hassane, D.C., Bonassar, L.J., Butcher, J.T., Weiss, R.S. & Reinhart-King, C.A. 2017. Matrix stiffening promotes a tumor vasculature phenotype. *Proceedings of the National Academy of Sciences of the United States of America* 114(3): 492-497.

Cai, F., Miao, Y., Liu, C., Wu, T., Shen, S., Su, X. & Shi, Y. 2018. Pyrroline-5-carboxylate reductase 1 promotes proliferation and inhibits apoptosis in non-small cell lung cancer. *Oncology Letters* 15(1): 731-740.

Cancer Genome Atlas Research Network. 2012. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* 489(7417): 519-525.

Castellano, B.M., Thelen, A.M., Moldavski, O., Feltes, M., Van Der Welle, R.E.N., Mydock-McGrane, L., Jiang, X., van Eijkeren, R.J., Davis, O.B., Louie, S.M., Perera, R.M., Covey, D.F., Nomura, D.K., Ory, D.S. & Zoncu, R. 2017. Lysosomal cholesterol activates mTORC1 via an SLC38A9–Niemann-Pick C1 signaling complex. *Science* 355(6331): 1306-1311.

Cau, F., Fanni, D., Manchia, M., Gerosa, C., Piras, M., Murru, R., Paribello, P., Congiu, T., Coni, P., Pichiri, G., Piludu, M., Van Eyken, P., Gibo, Y., La Nasa, G., Orrù, G., Scano, A., Coghe, F., Saba, L., Castagnola, M. & Faa, G. 2022. Expression of L1 cell adhesion molecule (L1CAM) in extracellular vesicles in the human spinal cord during development. *European Review for Medical and Pharmacological Sciences* 26(17): 6273-6282.

Cheng, H., Shcherba, M., Pendurti, G., Liang, Y., Piperdi, B. & Perez-Soler, R. 2014. Targeting the PI3K/AKT/mTOR pathway: Potential for lung cancer treatment. *Lung Cancer Management* 3(1): 67-75.

Choi, Y., Sims, G.E., Murphy, S., Miller, J.R. & Chan, A.P. 2012. Predicting the functional effect of amino acid substitutions and indels. *PLoS ONE* 7(10): e46688.

Christensen, E.M., Patel, S.M., Korasick, D.A., Campbell, A.C., Krause, K.L., Becker, D.F. & Tanner, J.J. 2017. Resolving the cofactor-binding site in the proline biosynthetic enzyme human pyrroline-5-carboxylate reductase 1. *Journal of Biological Chemistry* 292(17): 7233-7243.

Coradini, D., Ambrogi, F. & Infante, G. 2023. Cholesterol *de novo* biosynthesis in paired samples of breast cancer and adjacent histologically normal tissue: Association with proliferation index, tumor grade, and recurrence-free survival. *Archives of Breast Cancer* 10(2): 187-199.

Dennis, K.L., Blatner, N.R., Gounari, F. & Khazaie, K. 2013. Current status of IL-10 and regulatory T-cells in cancer. *Current Opinion in Oncology* 25(6): 637.

Ding, L., Getz, G., Wheeler, D.A., Mardis, E.R., McLellan, M.D., Cibulskis, K., Sougnez, C., Greulich, H., Muzny, D.M., Morgan, M.B., Fulton, L., Fulton, R.S., Zhang, Q., Wendl, M.C., Lawrence, M.S., Larson, D.E., Chen, K., Dooling, D.J., Sabo, A., Hawes, A.C., Shen, H., Jhangiani, S.N., Lewis, L.R., Hall, O., Zhu, Y., Mathew, T., Ren, Y., Yao, J., Scherer, S.E., Clerc, K., Metcalf, G.A., Ng, B., Milosavljevic, A., Gonzalez-Garay, M.L., Osborne, J.R., Meyer, R., Shi, X., Tang, Y., Koboldt, D.C., Lin, L., Abbott, R., Miner, T.L., Pohl, C., Fewell, G., Haipek, C., Schmidt, H., Dunford-Shore, B.H., Kraja, A., Crosby, S.D., Sawyer, C.S., Vickery, T., Sander, S., Robinson, J., Winckler, W., Baldwin, J., Chirieac, L.R., Dutt, A., Fennell, T., Hanna, M., Johnson, B.E., Onofrio, R.C., Thomas, R.K., Tonon, G., Weir, B.A., Zhao, X., Ziaugra, L., Zody, M.C., Giordano, T., Orringer, M.B., Roth, J.A., Spitz, M.R., Wistuba, I.I., Ozenberger, B., Good, P.J., Chang, A.C., Beer, D.G., Watson, M.A., Ladanyi, M., Broderick, S., Yoshizawa, A., Travis, W.D., Pao, W., Province, M.A., Weinstock, G.M., Varmus, H.E., Gabriel, S.B., Lander, E.S., Gibbs, R.A., Meyerson, M. & Wilson, R.K. 2008. Somatic mutations affect key pathways in lung adenocarcinoma. *Nature* 455(7216): 1069-1075.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T.R. 2013. STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* 29(1): 15-21.

Dou, X., Menkari, C., Mitsuyama, R., Foroud, T., Wetherill, L., Hammond, P., Suttie, M., Chen, X., Chen, S.Y., Charness, M.E. & Collaborative Initiative on Fetal Alcohol Spectrum Disorders. 2018. L1 coupling to ankyrin and the spectrin-actin cytoskeleton modulates ethanol inhibition of L1 adhesion and ethanol teratogenesis. *The FASEB Journal* 32(3): 1364-1374.

Dwyer-Nield, L.D., McArthur, D.G., Tennis, M.A., Merrick, D.T. & Keith, R.L. 2021. An improved murine premalignant squamous cell model: Tobacco smoke exposure augments NTCU-induced murine airway dysplasia. *Cancer Prevention Research* 14(3): 307-312.

Emmons, T. R., Giridharan, T., Singel, K. L., Khan, A. N. M. N. H., Ricciuti, J., Howard, K., Silva-Del Toro, S.L., Debreceni, I.L., Aarts, C.E.M., Brouwer, M.C., Suzuki, S., Kuijpers, T.W., Jongerius, I., Allen, L.H., Ferreira, V.P., Schubart, A., Sellner, H., Eder, J., Holland, S.M., Ram, S., Lederer, J.A., Eng, K.H., Moysich, K.B., Odunsi, K., Yaffe, M.B., Zsiros, E. & Segal, B.H. 2021. Mechanisms driving neutrophil-induced T-cell immunoparalysis in ovarian cancer. *Cancer Immunology Research* 9(7): 790-810.

Fang, B., Mehran, R.J., Heymach, J.V. & Swisher, S.G. 2015. Predictive biomarkers in precision medicine and drug development against lung cancer. *Chinese Journal of Cancer* 34(3): 1-15.

Fang, S., Dai, Y., Mei, Y., Yang, M., Hu, L., Yang, H., Guan, X. & Li, J. 2019. Clinical significance and biological role of cancer-derived Type I collagen in lung and esophageal cancers. *Thoracic Cancer* 10(2): 277-288.

Foretz, M., Pacot, C., Dugail, I., Lemarchand, P., Guichard, C., Le Lièpvre, X., Berthelier-Lubrano, C., Spiegelman, B., Kim, J.B., Ferré, P. & Foufelle, F. 1999. ADD1/SREBP-1c is required in the activation of hepatic lipogenic gene expression by glucose. *Molecular and Cellular Biology* 19(5): 3760-3768.

Gandara, D.R., Hammerman, P.S., Sos, M.L., Lara, P.N. & Hirsch, F.R. 2015. Squamous cell lung cancer: From tumor genomics to cancer therapeutics. *Clinical Cancer Research* 21(10): 2236-2243.

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): e0122823.

Gong, N., Wu, R., Ding, B. & Wu, W. 2020. ERBB4 promotes the progression of inflammatory breast cancer through regulating PDGFRA. *Translational Cancer Research* 9(5): 3266.

Gonzalez, H., Hagerling, C. & Werb, Z. 2018. Roles of the immune system in cancer: From tumor initiation to metastatic progression. *Genes & Development* 32(19-20): 1267-1284.

González, L.A., Melo-González, F., Sebastián, V.P., Vallejos, O.P., Noguera, L.P., Suazo, I.D., Schultz, B.M., Manosalva, A.H., Peñaloza, H.F., Soto, J.A., Parker, D., Riedel, C.A., González, P.A., Kalergis, A.M. & Bueno, S.M. 2021. Characterization of the anti-inflammatory capacity of IL-10-producing neutrophils in response to *Streptococcus pneumoniae* infection. *Frontiers in Immunology* 12: 638917.

Gu, L., Saha, S.T., Thomas, J. & Kaur, M. 2019. Targeting cellular cholesterol for anticancer therapy. *The FEBS Journal* 286(21): 4192-4208.

Gu, W., Zhuang, W., Zhuang, M., He, M. & Li, Z. 2023. DNA damage response and repair gene mutations are associated with tumor mutational burden and outcomes to platinum-based chemotherapy/immunotherapy in advanced NSCLC patients. *Diagnostic Pathology* 18(1): 119.

Guenzi, E., Pluvy, J., Guyard, A., Nguenang, M., Rebah, K., Benrahmoune, Z., Lamoril, J., Cazes, A., Gounant, V., Brosseau, S., Boileau, C., Zalcman, G. & Théou-Anton, N. 2021. A new *KIF5B–ERBB4* gene fusion in a lung adenocarcinoma patient. *ERJ Open Research* 7(1): 00582-2020.

Hamilton, J.A. & Achuthan, A. 2013. Colony stimulating factors and myeloid cell biology in health and disease. *Trends in Immunology* 34(2): 81-89.

Hartmann, P., Trufa, D.I., Hohenberger, K., Tausche, P., Trump, S., Mittler, S., Geppert, C.I., Rieker, R.J., Schieweck, O., Sirbu, H., Hartmann, A. & Finotto, S. 2023. Contribution of serum lipids and cholesterol cellular metabolism in lung cancer development and progression. *Scientific Reports* 13(1): 5662.

Henry, C.J., Billups, L.H., Avery, M.D., Rude, T.H., Dansie, D.R., Lopez, A., Sass, B., Whitmire, C.E. & Kouri, R.E. 1981. Lung cancer model system using 3-methylcholanthrene in inbred strains of mice. *Cancer Research* 41: 5027-5032.

Hoppstädter, J., Dembek, A., Höring, M., Schymik, H. S., Dahlem, C., Sultan, A., Wirth, N., Al-Fityan, S., Diesel, B., Gasparoni, G., Walter, J., Helms, V., Huwer, H., Simon, M., Liebisch, G., Schulz, M.H. & Kiemer, A.K. 2021. Dysregulation of cholesterol homeostasis in human lung cancer tissue and tumour-associated macrophages. *EBioMedicine* 72: 103578.

Hu, X., Xu, H., Xue, Q., Wen, R., Jiao, W. & Tian, K. 2021. The role of ERBB4 mutations in the prognosis of advanced non-small cell lung cancer treated with immune checkpoint inhibitors. *Molecular Medicine* 27: 126.

Huang, P., Nedelcu, D., Watanabe, M., Jao, C., Kim, Y., Liu, J. & Salic, A. 2016. Cellular cholesterol directly activates smoothened in hedgehog signaling. *Cell* 166(5): 1176-1187.

Hulsen, T., de Vlieg, J. & Alkema, W. 2008. BioVenn–A web application for the comparison and visualization of biological lists using area-proportional Venn diagrams. *BMC Genomics* 9(1): 488.

Janiszewska, M., Primi, M.C. & Izard, T. 2020. Cell adhesion in cancer: Beyond the migration of single cells. *Journal of Biological Chemistry* 295(8): 2495-2505.

Jiang, S., Wang, X., Song, D., Liu, X., Gu, Y., Xu, Z., Wang, X., Zhang, X., Ye, Q., Tong, Z., Yan, B., Yu, J., Chen, Y., Sun, M., Wang, Y. & Gao, S. 2019. Cholesterol induces epithelial-to-mesenchymal transition of prostate cancer cells by suppressing degradation of EGFR through APMAP. *Cancer Research* 79(12): 3063-3075.

Kawano, M., Mabuchi, S., Matsumoto, Y., Sasano, T., Takahashi, R., Kuroda, H., Kozasa, K., Hashimoto, K., Isobe, A., Sawada, K., Hamasaki, T., Morii, E. & Kimura, T. 2015. The significance of G-CSF expression and myeloid-derived suppressor cells in the chemoresistance of uterine cervical cancer. *Scientific Reports* 5(1): 18217.

Khoshnoodi, J., Sigmundsson, K., Öfverstedt, L-G., Skoglund, U., Öbrink, B., Wartiovaara, J. & Tryggvason, K. 2003. Nephrin promotes cell-cell adhesion through homophilic interactions. *The American Journal of Pathology* 163(6): 2337-2346.

Kim, S.H., Baek, S.I., Jung, J., Lee, E.S., Na, Y., Hwang, B.Y., Roh, Y.S., Hong, J.T., Han, S.B. & Kim, Y. 2022. Chemical inhibition of TRAF6-TAK1 axis as therapeutic strategy of endotoxin-induced liver disease. *Biomedicine & Pharmacotherapy* 155: 113688.

Kircher, M., Witten, D.M., Jain, P., O'roak, B.J., Cooper, G.M. & Shendure, J. 2014. A general framework for estimating the relative pathogenicity of human genetic variants. *Nature Genetics* 46(3): 310-315.

Korpanty, G.J., Graham, D.M., Vincent, M.D. & Leighl, N.B. 2014. Biomarkers that currently effect clinical practice in lung cancer: EGFR, ALK, MET, ROS-1 and KRAS. *Frontiers in Oncology* 4: 204.

Kourtidis, A., Lu, R., Pence, L.J. & Anastasiadis, P.Z. 2017. A central role for cadherin signaling in cancer. *Experimental Cell Research* 358(1): 78-85.

Kowanetz, M., Wu, X., Lee, J., Tan, M., Hagenbeek, T., Qu, X., Yu, L., Ross, J., Korsisaari, N., Cao, T., Bou-Reslan, H., Kallop, D., Weimer, R., Ludlam, M.J., Kaminker, J.S., Modrusan, Z., van Bruggen, N., Peale, F.V., Carano, R., Meng, Y.G. & Ferrara, N. 2010. Granulocyte-colony stimulating factor promotes lung metastasis through mobilization of Ly6G+ Ly6C+ granulocytes. *Proceedings of the National Academy of Sciences* 107(50): 21248-21255.

Kuzu, O.F., Gowda, R., Noory, M.A. & Robertson, G.P. 2017. Modulating cancer cell survival by targeting intracellular cholesterol transport. *British Journal of Cancer* 117(4): 513-524.

Lau, C., Killian, K.J., Samuels, Y. & Rudloff, U. 2014. ERBB4 mutation analysis: Emerging molecular target for melanoma treatment. *Molecular Diagnostics for Melanoma: Methods and Protocols* 1102: 461-480.

Lewis, C.A., Brault, C., Peck, B., Bensaad, K., Griffiths, B., Mitter, R., Chakravarty, P., East, P., Dankworth, B., Alibhai, D., Harris, A.L. & Schulze, A. 2015. SREBP maintains lipid biosynthesis and viability of cancer cells under lipid-and oxygen-deprived conditions and defines a gene signature associated with poor survival in glioblastoma multiforme. *Oncogene* 34(40): 5128-5140.

Li, J., Liu, N., Tang, L., Yan, B., Chen, X., Zhang, J. & Peng, C. 2020. The relationship between TRAF6 and tumors. *Cancer Cell International* 20: 429.

Liang, S., Xu, Y., Tan, F., Ding, L., Ma, Y. & Wang, M. 2018. Efficacy of icotinib in advanced lung squamous cell carcinoma. *Cancer Medicine* 7(9): 4456-4466.

Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P. & Mesirov, J. P. 2011. Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27(12): 1739-1740.

Liu, C.C., Lin, J.H., Hsu, T.W., Hsu, J.W., Chang, J.W., Su, K., Hsu, H.S. & Hung, S.C. 2018. Collagen XVII/laminin-5 activates epithelial-to-mesenchymal transition and is associated with poor prognosis in lung cancer. *Oncotarget* 9(2): 1656-1672.

Liu, L., Wu, Y., Zhang, C., Zhou, C., Li, Y., Zeng, Y., Zhang, C., Li, R., Luo, D., Wang, L., Zhang, L., Tu, S., Deng, H., Luo, S., Chen, Y.G., Xiong, X. & Yan, X. 2020. Cancer-associated adipocyte-derived G-CSF promotes breast cancer malignancy via Stat3 signaling. *Journal of Molecular Cell Biology* 12(9): 723-737.

Love, M.I., Huber, W. & Anders, S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* 15(12): 550.

Lucas, L.M., Dwivedi, V., Senfeld, J.I., Cullum, R.L., Mill, C.P., Piazza, J.T., Bryant, I.N., Cook, L.J., Miller, S.T., Lott 4th, J.H., Kelley, C.M., Knerr, E.L., Markham, J.A., Kaufmann, D.P., Jacobi, M.A., Shen, J. & Riese 2nd, D.J. 2022. The yin and yang of ERBB4: tumor suppressor and oncoprotein. *Pharmacological Reviews* 74(1): 18-47.

Masroor, M., Javid, J., Mir, R., Mohan, A., Ray, P.C. & Saxena, A. 2016. Prognostic significance of serum ERBB 3 and ERBB 4 mRNA in lung adenocarcinoma patients. *Tumor Biology* 37: 857-863.

Mayfosh, A.J., Baschuk, N. & Hulett, M.D. 2019. Leukocyte heparanase: A double-edged sword in tumor progression. *Frontiers in Oncology* 9: 331.

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., Garimella, K., Altshuler, D., Gabriel, S., Daly, M. & DePristo, M.A. 2010. The genome analysis toolkit: A mapreduce framework for analyzing next-generation DNA sequencing data. *Genome Research* 20(9): 1297-1303.

Missiaen, R., Mazzone, M. & Bergers, G. 2018. The reciprocal function and regulation of tumor vessels and immune cells offers new therapeutic opportunities in cancer. *Seminars in Cancer Biology* 52: 107-116.

Mok, E.H.K. & Lee, T.K.W. 2020. The pivotal role of the dysregulation of cholesterol homeostasis in cancer: Implications for therapeutic targets. *Cancers* 12(6): 1410.

Mollinedo, F. 2019. Neutrophil degranulation, plasticity, and cancer metastasis. *Trends in Immunology* 40(3): 228-242.

Mootha, V.K., Lindgren, C.M., Eriksson, K.F., Subramanian, A., Sihag, S., Lehar, J., Puigserver, P., Carlsson, E., Ridderstråle, M., Laurila, E., Houstis, N., Daly, M.J., Patterson, N., Mesirov, J.P., Golub, T.R., Tamayo, P., Spiegelman, B., Lander, E.S., Hirschhorn, J.N., Altshuler, D. & Groop, L.C. 2003. PGC-1α-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. *Nature Genetics* 34(3): 267-273.

Morris, K.T., Khan, H., Ahmad, A., Weston, L.L., Nofchissey, R.A., Pinchuk, I.V. & Beswick, E.J. 2014. G-CSF and G-CSFR are highly expressed in human gastric and colon cancers and promote carcinoma cell proliferation and migration. *British Journal of Cancer* 110(5): 1211-1220.

Morris, K.T., Castillo, E.F., Ray, A.L., Weston, L.L., Nofchissey, R.A., Hanson, J.A., Samedi, V.G., Pinchuk, I.V., Hudson, L.G. & Beswick, E.J. 2015. Anti-G-CSF treatment induces protective tumor immunity in mouse colon cancer by promoting NK cell, macrophage and T cell responses. *Oncotarget* 6(26): 22338.

Natarajan, S., Foreman, K.M., Soriano, M.I., Rossen, N.S., Shehade, H., Fregoso, D.R., Eggold, J.T., Krishnan, V., Dorigo, O., Krieg, A.J., Heilshorn, S.C., Sinha, S., Fuh, K.C. & Rankin, E.B. 2019. Collagen remodeling in the hypoxic tumor-mesothelial niche promotes ovarian cancer metastasis. *Cancer Research* 79(9): 2271-2284.

Okazaki, T., Ebihara, S., Asada, M., Kanda, A., Sasaki, H. & Yamaya, M. 2006. Granulocyte colony-stimulating factor promotes tumor angiogenesis via increasing circulating endothelial progenitor cells and Gr1+ CD11b+ cells in cancer animal models. *International Immunology* 18(1): 1-9.

Pan, J., Xiong, D., Zhang, Q., Szabo, E., Miller, M.S., Lubet, R.A., Wang, Y. & You, M. 2018. Airway brushing as a new experimental methodology to detect airway gene expression signatures in mouse lung squamous cell carcinoma. *Scientific Reports* 8(1): 8895.

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H. & Vilo, J. 2019. g: Profiler: A web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Research* 47(W1): W191-W198.

Riolobos, L., Gad, E.A., Treuting, P.M., Timms, A.E., Hershberg, E.A., Corulli, L.R., Rodmaker, E. & Disis, M.L. 2019. The effect of mouse strain, sex, and carcinogen dose on toxicity and the development of lung dysplasia and squamous cell carcinomas in mice. *Cancer Prevention Research* 12(8): 507-516.

Rodriguez, J.M., Maietta, P., Ezkurdia, I., Pietrelli, A., Wesselink, J.J., Lopez, G., Valencia, A. & Tress, M.L. 2013. APPRIS: Annotation of principal and alternative splice isoforms. *Nucleic Acids Research* 41(Database issue): D110-117.

Sang, S., Zhang, C. & Shan, J. 2019. Pyrroline-5-carboxylate reductase 1 accelerates the migration and invasion of nonsmall cell lung cancer *in vitro*. *Cancer Biotherapy and Radiopharmaceuticals* 34(6): 380-387.

She, Y., Mao, A., Li, F. & Wei, X. 2019. P5CR1 protein expression and the effect of gene-silencing on lung adenocarcinoma. *PeerJ* 7: e6934.

Shimano, H., Yahagi, N., Amemiya-Kudo, M., Hasty, A.H., Osuga, J., Tamura, Y., Shionoiri, F., Iizuka, Y., Ohashi, K., Harada, K., Gotoda, T., Ishibashi, S. & Yamada, N. 1999. Sterol regulatory element-binding protein-1 as a key transcription factor for nutritional induction of lipogenic enzyme genes. *Journal of Biological Chemistry* 274(50): 35832-35839.

Siegel, R.L., Giaquinto, A.N. & Jemal, A. 2024. Cancer statistics, 2024. *CA: A Cancer Journal for Clinicians* 74(1): 12-49.`

Strell, C., Lang, K., Niggemann, B., Zaenker, K.S. & Entschladen, F. 2010. Neutrophil granulocytes promote the migratory activity of MDA-MB-468 human breast carcinoma cells via ICAM-1. *Experimental Cell Research* 316(1): 138-148.

Subramanian, A., Tamayo, P., Mootha, V.K., Mukherjee, S., Ebert, B.L., Gillette, M.A., Paulovich, A., Pomeroy, S.L., Golub, T.R., Lander, E.S. & Mesirov, J.P. 2005. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proceedings of the National Academy of Sciences* 102(43): 15545-15550.

Sun, F., Yang, X., Jin, Y., Chen, L., Wang, L., Shi, M., Zhan, C., Shi, Y. & Wang, Q. 2017. Bioinformatics analyses of the differences between lung adenocarcinoma and squamous cell carcinoma using The Cancer Genome Atlas expression data. *Molecular Medicine Reports* 16(1): 609-616.

Surien, O., Ghazali, A.R. & Masre, S.F. 2020. Histopathological effect of pterostilbene as chemoprevention in N-nitroso-tri-chloroethylurea (NTCU)-Induced lung squamous cell carcinoma (SCC) mouse model. *Histology and Histopathology* 35(10): 1159-1170.

Suzuki, A., Onodera, K., Matsui, K., Seki, M., Esumi, H., Soga, T., Sugano, S., Kohno, T., Suzuki, Y. & Tsuchihara, K. 2019. Characterization of cancer omics and drug perturbations in panels of lung cancer cells. *Scientific Reports* 9(1): 19529.

Taghehchian, N., Moghbeli, M., Mashkani, B. & Abbaszadegan, M.R. 2021. The level of mesenchymal-epithelial transition autophosphorylation is correlated with esophageal squamous cell carcinoma migration. *Iranian Biomedical Journal* 25(4): 243.

Tamiya, M., Kobayashi, M., Morimura, O., Yasue, T., Nakasuji, T., Satomu, M., Kohei, O., Takayuki, S., Morishita, N., Suzuki, H., Sasada, S., Okamoto, N., Hirashima, T. & Kawase, I. 2013. Clinical significance of the serum crosslinked N-telopeptide of type i collagen as a prognostic marker for non-small-cell lung cancer. *Clinical Lung Cancer* 14(1): 50-54.

Tan, A.C. & Tan, D.S.W. 2022. Targeted therapies for lung cancer patients with oncogenic driver molecular alterations. *Journal of Clinical Oncology* 40(6): 611-625.

Tsai, J.H. & Yang, J. 2013. Epithelial–mesenchymal plasticity in carcinoma metastasis. *Genes & Development* 27(20): 2192-2206.

Valencia, K., Sainz, C., Bértolo, C., de Biurrun, G., Agorreta, J., Azpilikueta, A., Larrayoz, M., Bosco, G., Zandueta, C., Redrado, M., Redín, E., Exposito, F., Serrano, D., Echepare, M., Ajona, D., Melero, I., Pio, R., Thomas, R., Calvo, A. & Montuenga, L.M. 2022. Two cell line models to study multiorganic metastasis and immunotherapy in lung squamous cell carcinoma. *Disease Models & Mechanisms* 15(1): dmm049137.

Verstrepen, L., Bekaert, T., Chau, T.L., Tavernier, J., Chariot, A. & Beyaert, R. 2008. TLR-4, IL-1R and TNF-R signaling to NF-κB: Variations on a common theme. *Cellular and Molecular Life Sciences* 65: 2964-2978.

Voiles, L., Lewis, D.E., Han, L., Lupov, I.P., Lin, T.L., Robertson, M.J., Petrache, I. & Chang, H.C. 2014. Overexpression of type VI collagen in neoplastic lung tissues. *Oncology Reports* 32(5): 1897-1904.

Wang, B.Y., Huang, J.Y., Chen, H.C., Lin, C.H., Lin, S.H., Hung, W.H. & Cheng, Y.F. 2020. The comparison between adenocarcinoma and squamous cell carcinoma in lung cancer patients. *Journal of Cancer Research and Clinical Oncology* 146(3): 43-52.

Wang, D., Wang, L., Zhang, Y., Yan, Z., Liu, L. & Chen, G. 2019. PYCR1 promotes the progression of non-small-cell lung cancer under the negative regulation of miR-488. *Biomedicine & Pharmacotherapy* 111: 588-595.

Wang, Q.L. & Liu, L. 2019. PYCR1 is associated with papillary renal cell carcinoma progression. *Open Medicine* 14(1): 586-592.

Wang, Y., Zhang, Z., Garbow, J.R., Rowland, D.J., Lubet, R.A., Sit, D., Law, F. & You, M. 2009. Chemoprevention of lung squamous cell carcinoma in mice by a mixture of Chinese herbs. *Cancer Prevention Research* 2(7): 634-640.

Wang, Y., Zhang, Z., Yan, Y., Lemon, W.J., Laregina, M., Morrison, C. & Lubet, R. 2004. A chemically induced model for squamous cell carcinoma of the lung in mice : Histopathology and strain susceptibility. *Cancer Research* 64(5): 1647-1654.

Weijin, F., Zhibin, X., Shengfeng, Z., Xiaoli, Y., Qijian, D., Jiayi, L., Qiumei, L., Yilong, C., Hua, M., Deyun, L. & Jiwen, C. 2019. The clinical significance of PYCR1 expression in renal cell carcinoma. *Medicine* 98(28): e16384.

Williams, C.S., Bernard, J.K., Demory Beckler, M., Almohazey, D., Washington, M.K., Smith, J.J. & Frey, M.R. 2015. ERBB4 is over-expressed in human colon cancer and enhances cellular transformation. *Carcinogenesis* 36(7): 710-718.

Xiong, D., Pan, J., Yin, Y., Jiang, H., Szabo, E., Lubet, R.A., Wang, Y. & You, M. 2018. Novel mutational landscapes and expression signatures of lung squamous cell carcinoma. *Oncotarget* 9(7): 7424-7441.

Xu, S., Xu, H., Wang, W., Li, S., Li, H., Li, T., Zhang, W., Yu, X. & Liu, L. 2019. The role of collagen in cancer: From bench to bedside. *Journal of Translational Medicine* 17(1): 309.

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. doi:10.1111/cas.12855

Yi, M., Dong, Y., Zheng, M., Barr, M.P., Roviello, G., Zhihuang, H. & Liu, J. 2024. Identification of a prognostic gene signature in patients with cisplatin resistant squamous cell lung cancer. *Journal of Thoracic Disease* 16: 4567-4583.

Yoshimoto, T., Inoue, T., Iizuka, H., Nishikawa, H., Sakatani, M., Ogura, T., Hirao, F. & Yamamura, Y. 1980. Differential induction of squamous cell carcinomas and adenocarcinomas in mouse lung by intratracheal instillation of benzo(a)pyrene and charcoal powder. *Cancer Research* 40(11): 4301-4307.

Yoshimoto, T., Hirao, F., Sakatani, M., Nishikawa, H., Ogura, T. & Yamamura, Y. 1977. Induction of squamous cell carcinoma in the lung of C57BL/6 mice by intratracheal instillation of benzo [a] pyrene with charcoal powder. *GANN Japanese Journal of Cancer Research* 68(3): 343-352.

Yu, J.M., Sun, W., Hua, F., Xie, J., Lin, H., Zhou, D.D. & Hu, Z.W. 2015. BCL6 induces EMT by promoting the ZEB1-mediated transcription repression of E-cadherin in breast cancer cells. *Cancer Letters* 365(2): 190-200.

Zakaria, M.A., Aziz, J., Rajab, N.F., Chua, E.W. & Masre, S.F. 2022. Tissue rigidity increased during carcinogenesis of NTCU-induced lung squamous cell carcinoma *in vivo. Biomedicines* 10(10): 2382.

Zakaria, M.A., Rajab, N.F., Chua, E.W., Selvarajah, G.T. & Masre, S.F. 2021a. NTCU induced pre-malignant and malignant stages of lung squamous cell carcinoma in mice model. *Scientific Reports* 11(1): 22500.

Zakaria, M.A., Rajab, N.F., Chua, E.W., Selvarajah, G.T. & Masre, S.F. 2021b. Roles of Rho-associated kinase in lung cancer. *International Journal of Oncology* 58(2): 185-198.

Zeng, T., Zhu, L., Liao, M., Zhuo, W., Yang, S., Wu, W. & Wang, D. 2017. Knockdown of PYCR1 inhibits cell proliferation and colony formation via cell cycle arrest and apoptosis in prostate cancer. *Medical Oncology* 34(2): 27.

Zhang, X.C., Wang, J., Shao, G.G., Wang, Q., Qu, X., Wang, B., Moy, C., Fan, Y., Albertyn, Z., Huang, X., Zhang, J., Qiu, Y., Platero, S., Lorenzi, M.V., Zudaire, E., Yang, J., Cheng, Y., Xu, L. & Wu, Y.L. 2019. Comprehensive genomic and immunological characterization of Chinese non-small cell lung cancer patients. *Nature Communications* 10(1): 1772.

Zhang, Z., Wang, Y., Chen, L. & Li, Z. 2019. Protective effects of the suppressed NF-κB/TLR4 signaling pathway on oxidative stress of lung tissue in rat with acute lung injury. *The Kaohsiung Journal of Medical Sciences* 35(5): 265-276.

Zhao, W., Choi, Y. La, Song, J. Y., Zhu, Y., Xu, Q., Zhang, F., Jiang, L., Cheng, J., Zheng, G. & Mao, M. 2016. ALK, ROS1 and RET rearrangements in lung squamous cell carcinoma are very rare. *Lung Cancer* 94: 22-27.

Zilionis, R., Engblom, C., Pfirschke, C., Savova, V., Zemmour, D., Saatcioglu, H.D., Krishnan, I., Maroni, G., Meyerovitz, C.V., Kerwin, C.M., Choi, S., Richards, W.G., De Rienzo, A., Tenen, D.G., Bueno, R., Levantini, E., Pittet, M.J. & Klein, A.M. 2019. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* 50(5): 1317-1334.e.10.

*Corresponding author; email: sitifathiah@ukm.edu.my

SUPPLEMENTARY TABLE S1. The list of mutated genes unique in NTCU-induced LUSC in mice

ENSMUSG00000026247	ENSMUSG00000090588	ENSMUSG00000057072
ENSMUSG00000026385	ENSMUSG00000054951	ENSMUSG00000037318
ENSMUSG00000073530	ENSMUSG00000006678	ENSMUSG00000037892
ENSMUSG00000026941	ENSMUSG00000096768	ENSMUSG00000033233
ENSMUSG00000026754	ENSMUSG00000041263	ENSMUSG00000028602
ENSMUSG00000075307	ENSMUSG00000063779	ENSMUSG00000032661
ENSMUSG00000027004	ENSMUSG00000033308	ENSMUSG00000076548
ENSMUSG00000027224	ENSMUSG00000043633	ENSMUSG00000030325
ENSMUSG00000060336	ENSMUSG00000039577	ENSMUSG00000045467
ENSMUSG00000027748	ENSMUSG00000096351	ENSMUSG00000052273
ENSMUSG00000040896	ENSMUSG00000034462	ENSMUSG00000038457
ENSMUSG00000028134	ENSMUSG00000029672	ENSMUSG00000034218
ENSMUSG00000028782	ENSMUSG00000068335	ENSMUSG00000032419
ENSMUSG00000005907	ENSMUSG00000034783	ENSMUSG00000032500
ENSMUSG00000079451	ENSMUSG00000079494	ENSMUSG00000035448
ENSMUSG00000029516	ENSMUSG00000044086	ENSMUSG00000020330
ENSMUSG00000032850	ENSMUSG00000030321	ENSMUSG00000017550
ENSMUSG00000015942	ENSMUSG00000075589	ENSMUSG00000096466
ENSMUSG00000041453	ENSMUSG00000074417	ENSMUSG00000004341
ENSMUSG00000040187	ENSMUSG00000074158	ENSMUSG00000021640
ENSMUSG00000074361	ENSMUSG00000001773	ENSMUSG00000047911
ENSMUSG00000030484	ENSMUSG00000045087	ENSMUSG00000052496
ENSMUSG00000036862	ENSMUSG00000032036	ENSMUSG00000023806
ENSMUSG00000096401	ENSMUSG00000032087	ENSMUSG00000044043
ENSMUSG00000041624	ENSMUSG00000032355	
ENSMUSG00000057969	ENSMUSG00000032586	
ENSMUSG00000096054	ENSMUSG00000034579	
ENSMUSG00000020007	ENSMUSG00000055333	
ENSMUSG00000061315	ENSMUSG00000047904	
ENSMUSG00000075588	ENSMUSG00000025140	
ENSMUSG00000086022	ENSMUSG00000061603	
ENSMUSG00000004698	ENSMUSG00000072919	
ENSMUSG00000002799	ENSMUSG00000025876	
ENSMUSG00000056553	ENSMUSG00000040640	
ENSMUSG00000045410	ENSMUSG00000050335	
ENSMUSG00000035509	ENSMUSG00000057156	
ENSMUSG00000041707	ENSMUSG00000050463	
ENSMUSG00000046049	ENSMUSG00000044022	
ENSMUSG00000044309	ENSMUSG00000074882	
ENSMUSG00000071562	ENSMUSG00000051984	
ENSMUSG00000023868	ENSMUSG00000031398	
ENSMUSG00000024215	ENSMUSG00000096141	
ENSMUSG00000039512	ENSMUSG00000038209	

SUPPLEMENTARY TABLE S2. The list of ICGC donor ID with their respective treatment responsiveness class

ICGC donor ID	Treatment responsiveness class
DO26987	Treatment-responsive
DO26993	Treatment-responsive
DO26999	Treatment-responsive
DO26963	Treatment-responsive
DO26971	Non-responsive
DO26976	Treatment-responsive
DO26941	Treatment-responsive
DO26945	Non-responsive
DO26953	Non-responsive
DO26957	Treatment-responsive
DO26959	Treatment-responsive
DO26922	Non-responsive
DO26926	Non-responsive
DO26930	Non-responsive
DO26934	Treatment-responsive
DO26938	Treatment-responsive
DO26906	Treatment-responsive
DO26902	Treatment-responsive
DO26918	Non-responsive
DO26910	Non-responsive
DO26914	Treatment-responsive
DO26882	Treatment-responsive
DO26890	Treatment-responsive
DO26894	Non-responsive
DO26898	Non-responsive
DO26862	Treatment-responsive
DO26866	Non-responsive
DO26874	Treatment-responsive
DO26878	Non-responsive
DO26842	Treatment-responsive
DO26846	Non-responsive
DO26850	Non-responsive
DO26854	Non-responsive
DO26858	Non-responsive
DO26822	Treatment-responsive
DO26826	Treatment-responsive
DO26830	Treatment-responsive
DO26834	Non-responsive
DO26838	Non-responsive

continue to next page

continue from previous page

DO26806	Treatment-responsive
DO26818	Non-responsive
DO26810	Treatment-responsive
DO26520	Treatment-responsive
DO26522	Non-responsive
DO26524	Treatment-responsive
DO26526	Treatment-responsive
DO26528	Treatment-responsive
DO26530	Treatment-responsive
DO26532	Treatment-responsive
DO26534	Treatment-responsive
DO26536	Non-responsive
DO26538	Treatment-responsive
DO26502	Treatment-responsive
DO26504	Non-responsive
DO26508	Non-responsive
DO26510	Non-responsive
DO26512	Treatment-responsive
DO26514	Treatment-responsive
DO26516	Treatment-responsive
DO26518	Treatment-responsive
DO26590	Treatment-responsive
DO26584	Treatment-responsive
DO26587	Treatment-responsive
DO26593	Non-responsive
DO26596	Treatment-responsive
DO26599	Treatment-responsive
DO26560	Treatment-responsive
DO26563	Treatment-responsive
DO26566	Treatment-responsive
DO26569	Non-responsive
DO26581	Non-responsive
DO26572	Non-responsive
DO26575	Non-responsive
DO26578	Non-responsive
DO26540	Treatment-responsive
DO26542	Treatment-responsive
DO26544	Non-responsive
DO26546	Treatment-responsive
DO26548	Treatment-responsive

continue to next page

continue from previous page

DO26550	Non-responsive
DO26552	Non-responsive
DO26553	Non-responsive
DO26555	Non-responsive
DO26557	Non-responsive
DO27739	Non-responsive
DO27747	Treatment-responsive
DO27715	Non-responsive
DO27723	Treatment-responsive
DO27707	Non-responsive
DO26490	Treatment-responsive
DO26484	Treatment-responsive
DO26486	Non-responsive
DO26488	Treatment-responsive
DO26494	Treatment-responsive
DO26496	Non-responsive
DO50101	Non-responsive
DO50102	Treatment-responsive
DO50103	Treatment-responsive
DO50104	Treatment-responsive
DO26470	Treatment-responsive
DO26461	Non-responsive
DO26462	Treatment-responsive
DO26463	Treatment-responsive
DO26464	Non-responsive
DO26465	Treatment-responsive
DO26466	Non-responsive
DO26467	Non-responsive
DO26468	Non-responsive
DO26469	Non-responsive
DO26480	Non-responsive
DO26481	Non-responsive
DO26482	Treatment-responsive
DO26472	Treatment-responsive
DO26473	Non-responsive
DO26477	Non-responsive
DO26478	Non-responsive
DO26479	Treatment-responsive
DO26440	Treatment-responsive
DO26441	Non-responsive

continue to next page

continue from previous page

DO26442	Non-responsive
DO26443	Non-responsive
DO26444	Treatment-responsive
DO26445	Treatment-responsive
DO26446	Non-responsive
DO26447	Treatment-responsive
DO26448	Treatment-responsive
DO26449	Non-responsive
DO26460	Treatment-responsive
DO26453	Treatment-responsive
DO26454	Non-responsive
DO26455	Non-responsive
DO26456	Non-responsive
DO26457	Non-responsive
DO26459	Treatment-responsive
DO26420	Non-responsive
DO26421	Non-responsive
DO26422	Treatment-responsive
DO27755	Treatment-responsive
DO26423	Non-responsive
DO26424	Treatment-responsive
DO26425	Treatment-responsive
DO26426	Non-responsive
DO26427	Treatment-responsive
DO26428	Treatment-responsive
DO26429	Non-responsive
DO26431	Treatment-responsive
DO26432	Treatment-responsive
DO26434	Treatment-responsive
DO26435	Non-responsive
DO26436	Treatment-responsive
DO26437	Non-responsive
DO26438	Non-responsive
DO49084	Treatment-responsive
DO49085	Treatment-responsive
DO49083	Treatment-responsive
DO50097	Non-responsive
DO50098	Non-responsive
DO50094	Treatment-responsive
DO50095	Non-responsive
DO26761	Treatment-responsive

continue to next page

continue from previous page

DO26764	Treatment-responsive
DO26766	Non-responsive
DO26769	Non-responsive
DO26772	Treatment-responsive
DO26775	Non-responsive
DO26778	Non-responsive
DO26740	Treatment-responsive
DO26743	Treatment-responsive
DO26746	Non-responsive
DO26752	Non-responsive
DO26755	Treatment-responsive
DO26758	Treatment-responsive
DO26722	Treatment-responsive
DO26725	Non-responsive
DO26728	Non-responsive
DO26734	Treatment-responsive
DO26737	Treatment-responsive
DO26701	Treatment-responsive
DO26704	Treatment-responsive
DO26707	Non-responsive
DO26719	Treatment-responsive
DO26710	Treatment-responsive
DO26713	Treatment-responsive
DO26716	Non-responsive
DO26790	Non-responsive
DO26782	Treatment-responsive
DO26786	Non-responsive
DO26794	Non-responsive
DO26798	Treatment-responsive
DO26641	Non-responsive
DO26644	Treatment-responsive
DO26650	Treatment-responsive
DO26653	Non-responsive
DO26656	Treatment-responsive
DO26659	Non-responsive
DO26620	Non-responsive
DO26623	Non-responsive
DO26626	Treatment-responsive
DO26629	Non-responsive
DO26632	Non-responsive
DO26635	Non-responsive

continue to next page

continue from previous page

DO26638	Non-responsive
DO26602	Treatment-responsive
DO26605	Non-responsive
DO26608	Treatment-responsive
DO26611	Non-responsive
DO26614	Non-responsive
DO26617	Non-responsive
DO26686	Treatment-responsive
DO26689	Treatment-responsive
DO26692	Treatment-responsive
DO26695	Non-responsive
DO26698	Non-responsive
DO26662	Treatment-responsive
DO26665	Treatment-responsive
DO26668	Treatment-responsive
DO26680	Non-responsive
DO26671	Treatment-responsive
DO26674	Treatment-responsive
DO26677	Non-responsive
DO45499	Non-responsive
DO45495	Treatment-responsive
DO45496	Non-responsive
DO45498	Treatment-responsive
DO45491	Treatment-responsive
DO45492	Treatment-responsive
DO45494	Treatment-responsive
DO45488	Treatment-responsive
DO45489	Treatment-responsive
DO45485	Treatment-responsive
DO45482	Treatment-responsive
DO45490	Treatment-responsive
DO45479	Treatment-responsive
DO45473	Non-responsive
DO45476	Treatment-responsive
DO45471	Non-responsive
DO45467	Treatment-responsive
DO45469	Treatment-responsive
DO45462	Treatment-responsive
DO45463	Treatment-responsive
DO45465	Treatment-responsive
DO45460	Treatment-responsive

continue to next page

continue from previous page

DO45461	Treatment-responsive
DO45459	Treatment-responsive
DO45456	Treatment-responsive
DO45457	Treatment-responsive
DO45458	Treatment-responsive
DO45451	Treatment-responsive
DO45453	Treatment-responsive
DO45454	Treatment-responsive
DO45450	Treatment-responsive
DO45448	Treatment-responsive
DO45449	Treatment-responsive
DO45444	Treatment-responsive
DO45445	Non-responsive
DO45446	Treatment-responsive
DO45447	Treatment-responsive
DO45440	Treatment-responsive
DO45442	Treatment-responsive
DO45443	Treatment-responsive
DO45437	Treatment-responsive
DO45438	Treatment-responsive
DO45439	Non-responsive
DO45433	Treatment-responsive
DO45434	Treatment-responsive
DO45435	Treatment-responsive
DO45436	Non-responsive
DO45430	Treatment-responsive
DO45431	Treatment-responsive
DO45432	Non-responsive
DO45426	Treatment-responsive
DO45427	Treatment-responsive
DO45428	Non-responsive
DO45429	Treatment-responsive
DO45425	Treatment-responsive
DO45552	Treatment-responsive
DO45547	Treatment-responsive
DO45549	Treatment-responsive
DO45543	Treatment-responsive
DO45546	Treatment-responsive
DO45540	Treatment-responsive
DO45537	Non-responsive
DO45532	Treatment-responsive

continue to next page

continue from previous page

DO45535	Treatment-responsive
DO45530	Non-responsive
DO45527	Treatment-responsive
DO45523	Treatment-responsive
DO45520	Treatment-responsive
DO45517	Treatment-responsive
DO45510	Treatment-responsive
DO45513	Treatment-responsive
DO45503	Treatment-responsive
DO45506	Treatment-responsive
DO45500	Treatment-responsive
DO45509	Treatment-responsive
DO27461	Treatment-responsive
DO27453	Treatment-responsive
DO27468	Treatment-responsive
DO27437	Non-responsive
DO27445	Non-responsive
DO27413	Non-responsive
DO27421	Non-responsive
DO27429	Treatment-responsive
DO27405	Non-responsive
DO27397	Non-responsive
DO27381	Treatment-responsive
DO27389	Treatment-responsive
DO27357	Treatment-responsive
DO27373	Non-responsive
DO27365	Treatment-responsive
DO27334	Non-responsive
DO27341	Treatment-responsive
DO27349	Treatment-responsive
DO27310	Treatment-responsive
DO27318	Non-responsive
DO27302	Non-responsive
DO27279	Non-responsive
DO27295	Treatment-responsive
DO27255	Non-responsive
DO27271	Treatment-responsive
DO27240	Treatment-responsive
DO27232	Non-responsive
DO27247	Non-responsive
DO27612	Treatment-responsive

continue to next page

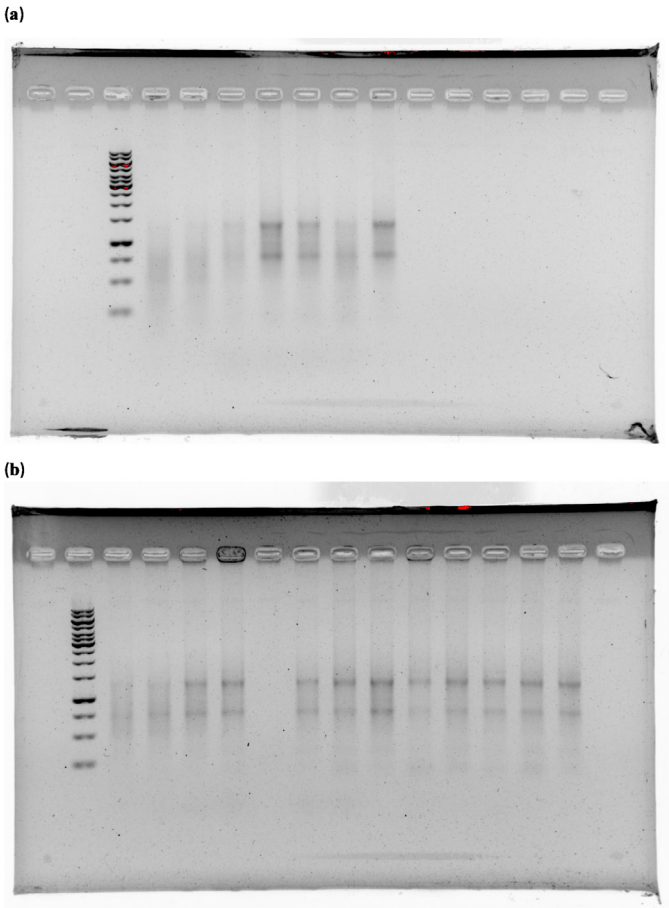
continue from previous page

DO27620	Treatment-responsive
DO27628	Treatment-responsive
DO27604	Non-responsive
DO27675	Treatment-responsive
DO27691	Treatment-responsive
DO27683	Non-responsive
DO27652	Treatment-responsive
DO27660	Non-responsive
DO27667	Treatment-responsive
DO27644	Treatment-responsive
DO27596	Treatment-responsive
DO27508	Treatment-responsive
DO27580	Non-responsive
DO27572	Treatment-responsive
DO27588	Non-responsive
DO27532	Treatment-responsive
DO27540	Non-responsive
DO27516	Treatment-responsive
DO27524	Non-responsive
DO27476	Non-responsive
DO27492	Treatment-responsive
DO27484	Non-responsive
DO27216	Treatment-responsive
DO27224	Treatment-responsive
DO27208	Non-responsive
DO27184	Treatment-responsive
DO27176	Non-responsive
DO27192	Treatment-responsive
DO27160	Non-responsive
DO27168	Treatment-responsive
DO27136	Treatment-responsive
DO27152	Treatment-responsive
DO27144	Non-responsive
DO27120	Non-responsive
DO27128	Treatment-responsive
DO27104	Treatment-responsive
DO52019	Treatment-responsive
DO27060	Non-responsive
DO52018	Treatment-responsive
DO27074	Treatment-responsive
DO27067	Non-responsive

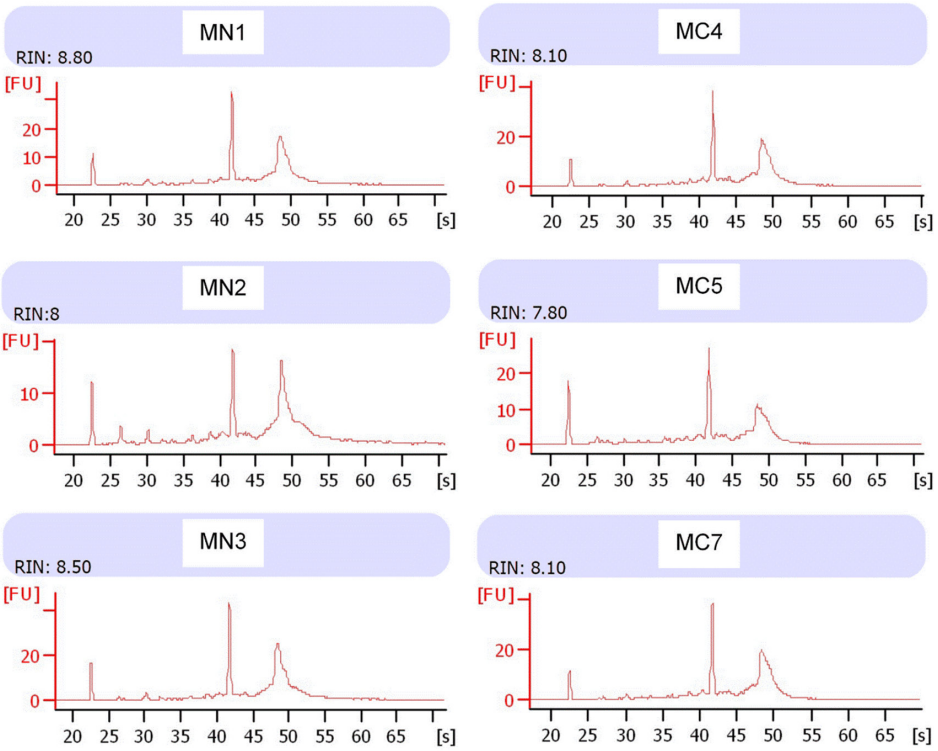
continue to next page

continue from previous page

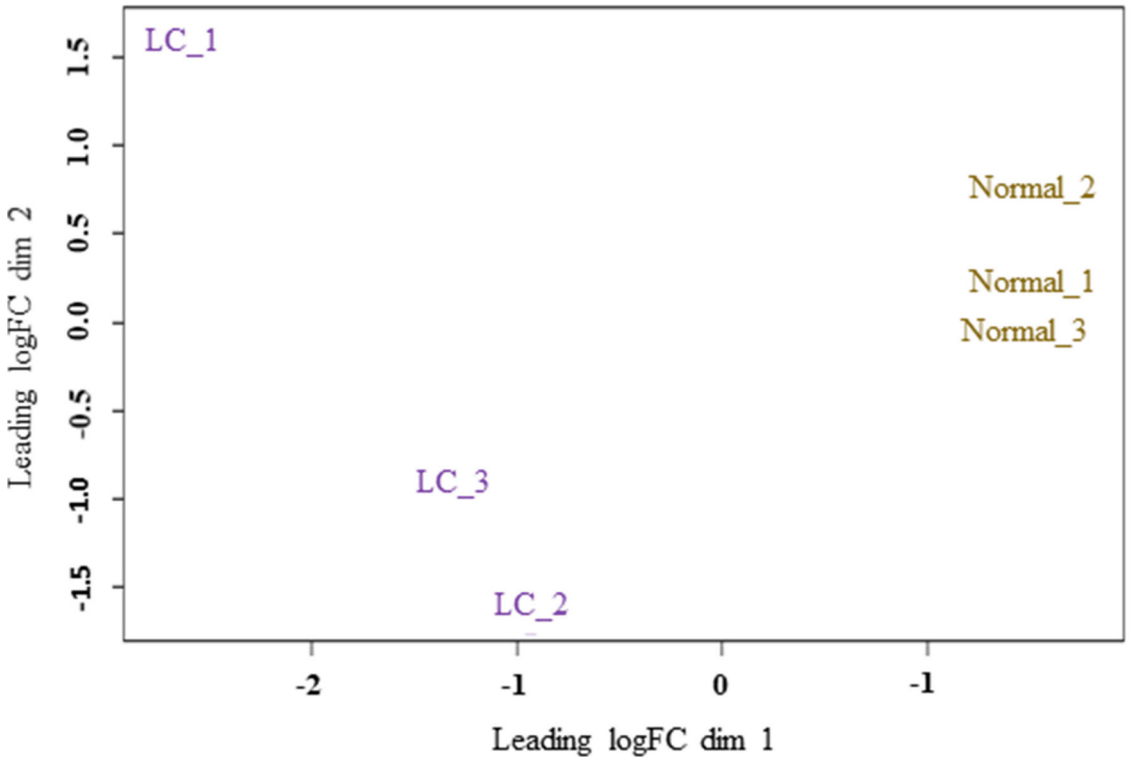
DO27039	Non-responsive
DO27053	Treatment-responsive
DO27046	Treatment-responsive
DO27011	Non-responsive
DO27018	Non-responsive
DO27032	Non-responsive
DO27025	Treatment-responsive
DO27005	Treatment-responsive
DO27081	Non-responsive
DO27096	Treatment-responsive
DO27088	Treatment-responsive
DO52023	Non-responsive
DO52025	Treatment-responsive
DO52026	Non-responsive
DO52020	Non-responsive
DO52021	Treatment-responsive



SUPPLEMENTARY FIGURE S1. (a) shows discrete 28s and 18s bands of RNA extracted from the vehicle groups and (b) shows discrete 28s and 18s bands of RNA extracted from the cancer groups.



SUPPLEMENTARY FIGURE S2. The RIN value of all mice lung tissue samples analyzed from the Vehicle (MC4, MC5 and MC7) and Cancer groups (MN1, MN2 and MN3).



SUPPLEMENTARY FIGURE S3. The PCA plot of all mice lung tissue samples analyzed from the Vehicle and Cancer groups.

1) The coding used in Anaconda prompt for in silico study (Phyton format) to convert the SSM data file in (.tsv) to text (.txt) format ready for VEP analysis.

```
#To get the file ready for VEP analysis.


import csv
import pandas as pd


with open('simple_somatic_mutation.open.tsv', 'r') as fin, open('VEP_input.txt', 'w') as fout:
        reader = csv.reader(fin, delimiter ='\t')
        writer = csv.writer(fout, delimiter ='\t')
        writer.writerow(['Chromosome', 'Chromosome_Start', 'Chromosome_End', 'Mutation'])
        next(reader, None) #To skip the headers.
        for row in reader:
                writer.writerow([row[8], row[9], row[10]] + ['/'.join([row[15], row[16]])])


df = pd.read_csv('VEP_input.txt', sep = '\t')
df.sort_values(by = ['Chromosome', 'Chromosome_Start'], inplace = True)
df.drop_duplicates(subset = ['Chromosome', 'Chromosome_Start', 'Chromosome_End', 'Mutation'],
inplace = True)
df.to_csv('VEP_input.txt', sep = '\t', index = False, header = False)
```

2) The coding used in Anaconda prompt for in silico study (Phyton format) to filter out non-major gene transcript and add VEP annotation (CADD scores) to the mutation file.

```
import csv
from itertools import islice


#To filter out non-major gene transcripts.


dominant = []


with open('appris_data.principal.txt','r') as fin:
        for row in csv.reader(fin, delimiter = '\t'):
                dominant.append(row[2])


with open('simple_somatic_mutation.open.tsv', 'r') as fin, open('ssm_major.txt', 'w', newline = '') as
fout:
        writer = csv.writer(fout, delimiter = '\t')
        for row in csv.reader(islice(fin, 1, None), delimiter ='\t'): #Using islice to skip one line.
                if row[29] in dominant:
                        writer.writerow([row[0], row[1], row[29]] + ['_'.join([row[8], row[9]] +
['/'.join([row[15], row[16]])])])


#To add VEP annotations to the mutation file.


with open('CADD_scores.txt', 'r') as fin:
        CADD_scores = {row[0]:row[41] for row in csv.reader(fin, delimiter = '\t')}


with open('ssm_major.txt', 'r') as fin, open('ssm_major_annotated.txt', 'w', newline = '') as fout:
        reader = csv.reader(fin, delimiter = '\t')
        writer = csv.writer(fout, delimiter = '\t')
        for row in reader:
                for key, value in CADD_scores.items():
                        if row[3] in key:
                                writer.writerow([row[0], row[1], row[2], row[3], value])
```

3) The coding used in Anaconda prompts for in silico study (Phyton format) to aggregate the list of pathogenic variants in an individual gene.

```
#To prepare data for downstream analysis e.g. g:Profiler.


import pandas as pd


df = pd.read_csv('ssm_major_annotated.txt', sep = '\t', names = ['Mutation_ID', 'Donor_ID',
'Major_Transcript', 'Mutation',
        'CADD_Score'])
df.drop(df[df.CADD_Score == '-'].index, inplace = True)
df['CADD_Score'] = pd.to_numeric(df['CADD_Score'])
df_gpd = df.groupby(by = ['Major_Transcript'], as_index = True)['CADD_Score'].sum()
df_gpd.to_csv('ssm_major_annotated_aggregated.txt', sep='\t', header = False)
```

SUPPLEMENTARY DATA S1: (Data S1)