Bilecik Şeyh Edebali Üniversitesi Fen Bilimleri Dergisi

Bilecik Seyh Edebali University Journal of Science

11(2), 276-284, 2024

DOİ: 10.35193/bseufbd.1361618

e-ISSN: 2458-7575 (https://dergipark.org.tr/tr/pub/bseufbd)

Araştırma Makalesi - Research Article

Gene Expression Profiling with Transcriptomic Data Analysis In Small Cell Lung Cancer

Küçük Hücreli Akciğer Kanserinde Transkriptomik Veri Analizi İle Gen Ekspresyon Profili

Gözde Öztan[1*]

Geliş / Received: 16/09/2023 ***Revize / Revised: 03/12/2023*** ***Kabul / Accepted: 17/12/2023***

ABSTRACT

Small-cell lung cancer (SCLC) is aggressive due to fast tumor development, early metastatic dissemination, and genetic instability. In this study, the RNA sequencing method was applied to the selected experimental data set for gene expression analysis in lung tissue samples of SCLC using Array Express functional genomic data. Array Express is a public repository for transcriptomic and related data that aims to store MIAME-compliant data in accordance with MGED recommendations. We wanted to look into the genomic sequence data (GSE60052) of 7 healthy controls and 75 SCLC patients through the GEO2R platform and the NCBI Gene Expression Omnibus (GEO) using the accession number E-GEOD-60052. The GSE60052 dataset of the genomic expression study was found on the GEO2R platform using the Illumina HiSeq 2000 RNA sequencing method in lung tissue samples from 75 SCLC patients and 7 controls. This was done to find out how the gene profile in SCLC were being expressed. In patients both in the SCLC and the control group, it was identified through the Volcano plot graph that HOXD10, FAM83A, HOXB1, ECEL1, GATA4, DMRT3, TGM3, CHP2, and PPP1R1A genes were down-regulated (log2(fold change) < -5), while PGC, SFTPC, SLC6A4, and CSF3 genes were up-regulated (log2 (fold change > +5). We share the view that SCLC is a type of neuroendocrine tumor with high malignancy and a poor prognosis, and identifying significant genes through expression profiling in lung tissue samples may be effective in elucidating the complex mechanisms underlying SCLC and determining their effect on the prognosis of the disease. The use of related genes as possible prognostic biomarkers in targeted therapy in SCLC could be enables the determination of the effects of the tumor microenvironment on immune cells and stromal cells.

Keywords- Biomarker*, *Gene Expression, RNA Sequencing, Small Cell Lung Cancer, Transcription Profile Array

ÖZ

Küçük hücreli akciğer kanseri (SCLC), hızlı tümör gelişimi, erken metastatik yayılım ve genetik dengesizlik nedeniyle agresiftir. Array Express, MIAME uyumlu verileri MGED önerilerine uygun olarak depolamayı amaçlayan, transkriptomik ve ilgili veriler için halka açık bir depodur. Bu çalışmada, Array Express fonksiyonel genomik datası kullanarak, SCLC'nin akciğer dokusu örneklerinde gen ekspresyon analizine yönelik seçilen deneysel veri setinde RNA dizileme yöntemi uygulanmıştır. E-GEOD-60052 erişim numarası üzerinden NCBI Gene Ekspresyon Omnibus (GEO) kullanarak 7 sağlıklı kontrol ve 75 SCLC'li hastaların genomik dizi verilerini (GSE60052) GEO2R platformu aracılığıyla araştırmayı amaçladık. SCLC'de genlerin ekspresyon düzeylerinin

[1*]Corresponding Author Contact: gozdeoztan@istanbul.edu.tr (https://orcid.org/0000-0002-2970-1834)
Department of Medical Biology, Istanbul Faculty of Medicine, Istanbul University, Istanbul, Turkey

belirlenmesi için 75 SCLC hastasından ve 7 kontrolden alınan akciğer dokusu örneklerinde Illumina HiSeq 2000 RNA dizileme yöntemi kullanılarak genomik ekspresyon çalışmasının GSE60052 veri seti GEO2R platformu üzerinden tespit edildi. SCLC'li hastalar ve kontrol grubunda, HOXD10, FAM83A, HOXB1, ECEL1, GATA4, DMRT3, TGM3, CHP2, PPP1R1A genlerinin aşağı regüle olduğu (log2(fold change) < -5), PGC, SFTPC, SLC6A4, CSF3 genlerinin (log2 (fold change > +5) ise yukarı regüle olduğu Volcano plot grafiği aracılığıyla tanımlandı. Yüksek maligniteye ve kötü prognoza sahip bir tür nöroendokrin tümör olan SCLC'nin, akciğer dokusu örneklerinde ekspresyon profillemesiyle anlamlı bulunan genlerin tanımlanmasının SCLC'nin altında yatan karmaşık mekanizmalarının aydınlatılmasında ve hastalığın prognozuna olan etkisinin belirlenmesinde etkili olabileceği görüşünü paylaşmaktayız. İlgili genlerin SCLC'de hedefe yönelik tedavide olası prognostik biyobelirteçler olarak kullanılması, tümör mikro ortamının bağışıklık hücreleri ve stromal hücreler üzerindeki etkilerinin belirlenmesini sağlar.

Anahtar Kelimeler- Küçük Hücreli Akciğer Kanseri, Transkripsiyon Profil Array, Gen Ekspresyonu

I. INTRODUCTION

Lung cancer is the leading cause of cancer-related mortality worldwide [1]. Estimates from GLOBOCAN 2020 indicate that 2.2 million new cases of lung cancer (11.4%) and about 1.8 million deaths from lung cancer (18.0%) took place in 2020 [2]. Lung neoplasms rank second among cancer diagnoses in both men and women, behind only breast and prostate cancers, respectively [3]. Lung cancer kills 350 people daily, about 2.5 times more than colorectal cancer, the second-greatest cause of cancer deaths [4]. Small-cell lung cancer (SCLC) and non-small-cell lung cancer (NSCLC) are the two main subtypes of lung cancer that may be distinguished by histopathology. The most lethal form of lung cancer, SCLC, is responsible for 13–15% of all cases and has a 5-year survival rate of fewer than 7%. The course of SCLC is defined as being very aggressive due to an unstable tumor genome, a quick growth rate, enhanced angiogenesis, and a high propensity for metastatic spread [1]. Only about a third of individuals get their cancer detected at an early enough stage to benefit from multimodal treatment, while the majority have advanced metastatic disease. Genomic analysis of SCLC shows widespread chromosomal rearrangements and a significant mutation load, which almost always includes functional inactivation of TP53 and RB1. Subtypes of human SCLC have been established by analyzing the expression of prominent transcriptional regulators, and analyses of murine models have demonstrated significant intratumoral variability [5].

The embryonic development of several organs is controlled by the hedgehog signaling pathway. Adult stem cell renewal and organ homeostasis are both regulated by the hedgehog signaling pathway. Within the classical hedgehog signaling pathway, three ligands have been identified: Sonic Hedgehog, Indian Hedgehog, and Desert Hedgehog [6]. SCLC cells stimulate the hedgehog signaling pathway, which regulates airway epithelial shape and stem-cell destiny throughout embryonic development. It is essential to activate the pathway in order to maintain the viability of SCLC cells both in vitro and in vivo [7]. One important route for physiological cell death in vertebrates is the mitochondrial pathway to apoptosis.The mitochondrial outer membrane permeabilization (MOMP) process starts the cell death pathway by letting apoptogenic chemicals into the cytosol from the inner and outer mitochondrial membranes that are intertwined. The BCL-2 proteins that help cells die (BCL-2 associated x protein and BCL-2 antagonist killer 1) start MOMP. The BCL-2 proteins that stop cells from dying (BCL-2, BCL-xl, and myeloid cell leukemia 1) stop it. Caspases are activated when mitochondria produce pro-apoptotic substances, including cytochrome c, which causes the creation of the apoptosome, a multimeric complex. While these pathways are critical for healthy cellular homeostasis, they are also involved in the development of many illnesses [8]. The mitochondrial apoptosis pathway is now being investigated as a potential therapy option for SCLC. There are proapoptotic and antiapoptotic functions for proteins in the Bcl-2 family, which are essential regulators of apoptosis. SCLC cell lines and primary tissues overexpress Bcl-2, which inhibits BAX and BAK [7].

Single-cell RNA sequencing, or scRNA-seq, has been used to study and describe the molecular features of lung cancer cells and the tumor microenvironment in great detail. Recent advances in high-throughput sequencing methods, particularly single-cell sequencing, have made it feasible to characterize lung cancer molecularly in great detail [9]. New generation single-cell sequencing method have made it possible to create a detailed profile of individual cells in the tumor microenvironment (TME) and their potential involvement in tumorigenic processes. Tumor cells in SCLCs showed a wide range of characteristics, most of which were associated with the cell cycle, the immune system, and hypoxia. The results obtained from the previous study showed that major transcription factors in SCLC exhibit intratumor variability in gene expression patterns and associated activities. A higher response to immune control inhibitors has been shown in SCLC, a non-neuroendocrine tumor, where there is a link between elevated inflammatory gene markers and immune cell infiltrates [10]. SCLC contains several chromosomal rearrangements and mutations, according to omics profiling [11]. The discovery of immune checkpoint inhibitors, either alone or in combination with chemotherapy, has

shown long-term benefits for only a tiny proportion of patients [9]. Recent advances in single-cell harvesting, liquid biopsies, and genomics-bioinformatics analysis hold promise as potent new methods for studying drug resistance. Still, there may be scientific grounds and a better clinical result for SCLC if we learn more about its features, such as tumor immunity, immunological microenvironment, intratumoral heterogeneity, and genetic profiles and development [12].

The scRNA-seq approach used in SCLC offers a thorough cellular diversity setting in lung tumor tissues, according to our study. The ability to compare and contrast tumor cell heterogeneity and cellular variations is therefore realized. Additionally, the scRNA-seq approach may reveal genetic processes that may lead to drug resistance after tumor therapy. Is there a relationship between important SCLC-related genes and the prognosis of the disease when creating an expression profile from SCLC lung tissue samples? The goal is to learn all the genes that are involved in SCLC cancer cell biology and phenotype by studying functional genomic data sets. Also, we discovered genes that could be used as biomarkers for SCLC that can help guide targeted treatment. The study's hypothesis can help us learn more about the molecular processes that lead to lung cancer. The aim of this study was to examine the underlying pathogenesis of SCLC and potential molecular markers and gene expression levels between patients with SCLC and healthy control groups.

II. MATERIAL AND METHOD

A. Selection of an experimental dataset via the Array Express platform

The primary objective of the first step of our investigation was to show gene expression patterns. To do this, we searched the Array Express database, which openly publishes high-throughput functional genomic data, utilizing gene names and characteristics such as Gene Ontology keywords. The Array Express database allows for the querying of all experiments that have disease expression, and it also allows for the accessing of data based on the accession numbers and sequence design names of the experiments, respectively. Array Express, which provides data analysis, mining, and visualization by providing easy access to large volumes of data, such as the transcriptomic data analysis we included in our study, via the web, has many filtering options such as study type, experimental design, organism, technology, assay by molecule, access to raw data, access to processed data, released, link type, and file type in order to provide deeper integration. Identification of differentially expressed genes by transcriptome sequencing (RNA array) will provide guidance in the pathogenesis of SCLC. Array Express is an open database for transcriptomic data that follows the MGED-recommended format for storing MIAME-compliant information (https://www.ebi.ac.uk/biostudies/arrayexpress). In this study, the experimental data set with the accession number E-GEOD-60052 was chosen. Samples of lung tissue from SCLC patients were sequenced using the Illumina HiSeq 2000 RNA sequencing technology and analyzed on the Array Express platform to determine gene expression.

B. Identification of non-coding RNAs via the NCBI Gene Expression Omnibus (GEO) database

Gene chip data were searched using the GEO database, a publicly available genomic database containing all gene expression data, chips, and microarrays. (https://www.ncbi.nlm.nih.gov/geo). The coding RNA sequence dataset (GSE60052) was downloaded from the GEO database. After that, we used the GEO database, which stores both raw and processed data related to high-throughput gene expression and genomic studies, including sample designs, methods, and experimental designs. As the GEO database continues to evolve, it will include an increasing amount of gene expression data. This opens up new possibilities for solving data complexity and conducting data queries, visualizations, and analyses straight from the GEO website. Filling up the "GEO access" query box with the GEO accession number chosen in the Array Express database does this. The next step is to run the data via the GEO2R program for analysis.

RNA sequencing data from 75 individuals diagnosed with SCLC and 7 healthy individuals was detected using the GEO2R program through GEO, a functional genomic data repository for the dataset with accession number E-GEOD-60052. GEO2R is a web-based, user-interactive application for analyzing expression profile array datasets for differentially expressed genes (DEGs). GEO2R's define group command makes it easy to categorize subjects into different categories. Simple analysis customization is made possible using GEO2R's panel choices. Selecting the statistical trimmer, data normalization technique, and cutoff value from the option panel enables users to exclude genes that do not meet the set cutoff value [13].

The "GEO2R," "Options," "Profile Plot," and "R Script" tabs are part of the GEO2R suite of tools. This GEO2R program is based on the R packages Limma and GEOQuery, both developed by Bioconductor. The Benjamini-Hochberg false discovery rate approach is used by default for multiple testing corrections. Limma and DESeq2 offer a variety of options for adjusting P-values. In an effort to account for the possibility of false positive findings, these modifications are known as multiple-testing corrections. By default, the Benjamini & Hochberg false discovery rate approach is used because it effectively limits false positives while also discovering statistically important genes.

C. Evaluation of gene data through the Human Protein Atlas database

We used the Human Protein Atlas database to sort the proteins made by the genes we found to be statistically significant using the RNA-Seq method into groups based on the type of cell they were in and to find out where in the body they were found. It can be a guide in determining cell-based treatment options specific to disease-related genes through the Human Protein Atlas database.

The Human Protein Atlas database's section on tissue cell type offers cell type expression specificity estimates for all human protein-coding genes. These estimates were developed by an integrated network analysis of publicly accessible bulk RNAseq data (https://www.proteinatlas.org). Estimating which genes were enriched in which cell types within a particular tissue was done using a specificity classification. Tissue slices stained with immunohistochemistry were used for the analysis. The cell-type specificity of genes in lung tissue was investigated. The cell types of the down-regulated and up-regulated genes in the lung tissue were determined, as were the enrichment scores of the genes.

III. RESULTS

The raw data of GSE60052 was efficiently processed using the limma package in R using smoothing, normalization, and log2 transformation. Using the Illumina HiSeq 2000 RNA sequencing technology in the GEO2R software, lung tissue samples from 75 SCLC patients and 7 controls were sequenced to find genes and pathways that are unique to SCLC. The GSE60052 dataset was from a genome-wide expression study.

The findings of RNA-seq or other omics investigations are often presented using volcano plot graphics. The genes that are most significantly up-regulated are shown to the right of a volcano plot, whereas the genes that are most significantly down-regulated are displayed to the left of the plot, and the most statistically significant genes are displayed at the top of the plot. Non-coding RNAs expressed at different levels are shown according to the log2 fold change in the volcano plot. Genes with a p-value of 0.05 are considered to be substantially differentially expressed.

According to the Volcano plot, homeobox D10 (HOXD10), family with sequence similarity 83 member A (FAM83A), homeobox B1 (HOXB1), endothelin converting enzyme like 1 (ECEL1), GATA binding protein 4 (GATA4), doublesex and mab-3 related transcription factor 3 (DMRT3), transglutaminase 3 (TGM3), calcineurin like EF-hand protein 2 (CHP2), and protein phosphatase 1 regulatory inhibitor subunit 1A (PPP1R1A) genes were downregulated (log2(fold change) <-5) (Table 1), while progastricsin (PGC), surfactant protein C (SFTPC), solute carrier family 6 member 4 (SLC6A4), and colony stimulating factor 3 (CSF3) genes (log2 (fold change > +5) were identified as upregulated (Figure 1) (Table 2).

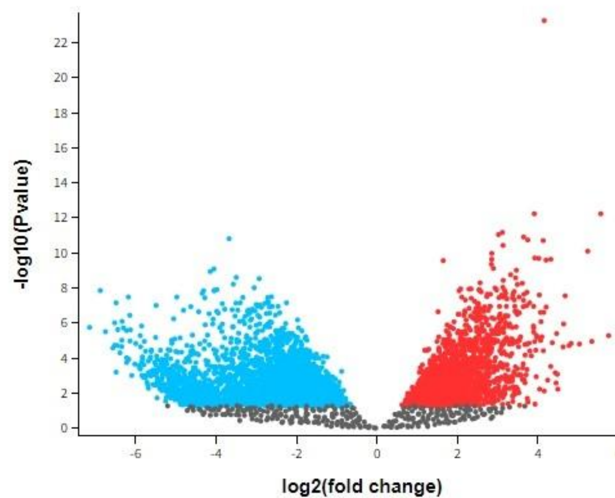

Figure 1. Volcano plot (comparison of patients with SCLC and a healthy control group)

Table 1. Down-regulated genes obtained in comparison of patients with SCLC and individuals in the healthy control group (log2(fold change) < -5)

Gene ID	Genes	log2 (fold change)	-log10 (p value)
3236	HOXD10	-6.177	7.488
84985	FAM83A	-6.333	6.099
3211	HOXB1	-6.386	4.305
9427	ECEL1	-5.124	3.392
2626	GATA4	-5.304	2.518
58524	DMRT3	-5.627	3.206
7053	TGM3	-5.091	2.392
63928	CHP2	-5.083	2.423
5502	PPP1R1A	-5.435	2.49

Table 2. Up-regulated genes were obtained in comparison between patients with SCLC and individuals in the healthy control group (log2 (fold change > +5)

Gene ID	Genes	log2 (fold change)	-log10 (p value)
5225	PGC	5.235	10.101
6440	SFTPC	5.337	4.956
6532	SLC6A4	5.559	12.235
1440	CSF3	5.029	4.809

By looking at the boxplot distribution, it was decided that some samples from people with SCLC and some samples from people in the healthy control group could be used for differential expression analysis. Expression values were determined using the limma package in R software (Figure 2). Black bars give the median value. The resulting median-centered values indicate that the data are normalized and cross-comparable.

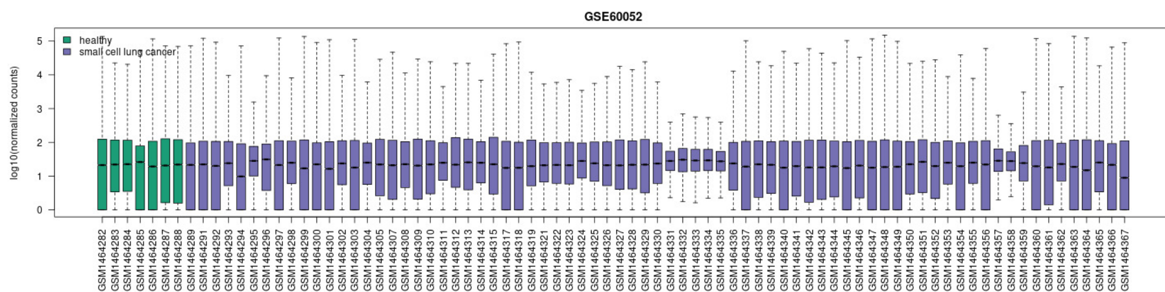

Figure 2. R boxplot displaying the distribution of sampled values

In the next part of the study, the Human Protein Atlas database (https://www.proteinatlas.org) was used to look at the genome-wide analysis of genes found to be important in SCLC that make proteins. Accordingly, among the down-regulated genes, the expression specificities of the HOXD10 gene in fibroblasts and mast cells, the FAM83A, HOXB1, TGM3, and CHP2 genes in respiratory ciliary cells, the ECEL1, GATA4, and DMRT3 genes in fibroblasts, and the PPP1R1A gene in smooth muscle cells were determined (Table 3).

Table 3. Cell types and enrichment scores in lung tissue of downregulated genes according to the Human Protein Atlas database

Genes	Cell type	Enrichment score
HOXD10	Fibroblast (mesenchymal cells)	0.346
HOXD10	Mast cells (blood and immune cells)	0.302
FAM83A	Respiratory ciliary cells (Glandular epithelial cells)	0.564
HOXB1	Respiratory ciliary cells (Glandular epithelial cells)	0.376
ECEL1	Fibroblast (mesenchymal cells)	0.442
GATA4	Fibroblast (mesenchymal cells)	0.318
DMRT3	Fibroblast (mesenchymal cells)	0.303
TGM3	Respiratory ciliary cells (Glandular epithelial cells)	0.596
CHP2	Respiratory ciliary cells (Glandular epithelial cells)	0.412
PPP1R1A	Smooth muscle cells (muscle cells)	0.443

Based on the Human Protein Atlas database, the expression specificities of the PGC gene in type II alveolar cells and macrophages, the SFTPC gene in type II alveolar cells and macrophages, the SLC6A4 gene in type I alveolar cells, and the CSF3 gene in neutrophils were found (Table 4).

Table 4. Cell types and enrichment scores in lung tissue of upregulated genes according to the Human Protein Atlas database

Genes	Cell type	Enrichment score
PGC	Type II alveolar cells (specialized epithelial cells)	0.748
PGC	Macrophages (blood and immune cells)	0.585
SFTPC	Type II alveolar cells (specialized epithelial cells)	0.830
SFTPC	Macrophages (blood and immune cells)	0.581
SLC6A4	Type I alveolar cells (specialized epithelial cells)	0.343
CSF3	Neutrophils (blood and immune cells)	0.302

IV. DISCUSSION

A complete genomic or proteomic study is essential to identifying cancer biology. The vast genetic study of many solid tumors, including SCLC, has only recently been possible because of advances in molecular biology. Next-generation sequencing aided in the genomic and epigenetic analyses; RNA microarrays helped identify the transcriptome; and mass spectrometry provided the proteomics [14]. Using omics data, researchers found that SCLC had a high mutation load and widespread chromosomal rearrangements.

Immune checkpoint inhibitors were developed as standalone medicines or in conjunction with chemotherapy, but only a tiny proportion of patients had a persistent benefit [11]. Despite an encouraging early response to systemic therapy, SCLC continues to be a malignancy with a dismal prognosis. The majority of patients eventually acquire severe disease, quickly develop resistance to therapy, and ultimately pass away as a result of

their illness. Patients with NSCLC have greatly benefited from targeted therapy, while those with SCLC have not seen the same progress. The progress made in our molecular and genetic knowledge of tumor biology in recent years, however, bodes well for big advances for patients with SCLC in the near future. PARP1, BCL-2, WEE1, EZH2, and DLL3 are only a few of the new targets in SCLC that have been uncovered by thorough genomic, proteomic, and transcriptome investigations [15].

HOXD10 belongs to the homeobox gene family and is a master regulator that can directly affect organogenesis and play a role in maintaining differentiated tissue functions [16]. They found the HOXD10 gene, which we found to be down-regulated in our study, to be up-regulated in a study conducted for the determination of biomarkers and pathways in SCLC [17]. High expression of FAM83A in lung adenocarcinoma (LUAD), the most common subtype of non-small cell lung cancer, predicted a poor prognosis in LUAD patients. Regardless of gender, low-expression individuals have a better prognosis than high-expression patients [18]. Through analysis of RNA-Seq data from the encyclopedia lung cancer cell line, the level of FAM83A antisense RNA 1 (FAM83A-AS1) was found to be higher in LUAD, squamous cell lung cancer (LUSC), and large cell lung cancer (LLC), but not in SCLC. It has been confirmed that it is not higher [19].

Specific hsa-let-7g effects on HOXB1 were reversed when HOXB1 expression was reduced in lung cancer cells. Hsa-let-7g overexpression has been linked to lung cancer progression by inhibiting HOXB1 expression [20]. Reduced GATA4 levels were strongly linked with a poor prognosis, and these levels were associated adversely with WNT7B or TGF-2. Recent research demonstrates the tumor-suppressing role of GATA4 in lung cancer and suggests that therapeutic targeting of TGF-β signaling may be an effective strategy for treating GATA4-deficient lung cancer [21]. In our study, a parallel decrease in GATA4 expression level was observed, and we can say that the course of the disease is associated with a poor prognosis in the SCLC patients we selected in the study.

DMRT3 was found to be significantly upregulated in LUAD cell lines. High DMRT3 expression has been observed to be significantly associated with poor overall survival in LUAD [22]. The copy number of DMRT3 is different in different LUSC tumors. Tumors that don’t have DMRT3 deletions express DMRT3 much more than tumors that do [23]. We determined that the DMRT3 gene was downregulated in the SCLC patient group. Compared with healthy control groups, TGM3 expression was found to be higher in cancer groups, including LUAD and LUSC [24]. We identified the TGM3 gene as downregulated in the patient group with SCLC.

New evidence suggests that the protein CHP2 promotes NSCLC tumor growth, suggesting that it might be used as a prognostic indicator or therapeutic target for the illness. Upregulation of CHP2 was seen in NSCLC tissues and cells, and data showed that high CHP2 levels were inversely related to patients’ 5-year survival rates [25]. In our study, we found that the expression level of the CHP2 gene was downregulated in patients with SCLC. The expression of the PPP1R1A gene has been detected in normal tissues and has also been determined to be reduced in various lung cancer cell lines [26]. Based on these results, it can be said that similar results were obtained with our study. When we compared patients with SCLC to healthy controls, we determined that four genes: PGC, SFTPC, SLC6A4, and CSF3 were upregulated. One study reported that circular progastricsin (circ-PGC) was upregulated in NSCLC tissues and cells. Circ-PGC knockdown has been determined to inhibit NSCLC cell viability, colony formation, cell migration, invasion, and glycolysis metabolism [27].

The SFTPC gene encodes a protein known as pulmonary-associated surfactant protein C, which is necessary for lung function and homeostasis. Low levels of SFTPC expression were linked to poor overall survival in lung adenocarcinoma patients, and both human lung cancer tissues and cell lines showed downregulation of SFTPC expression. SFTPC overexpression inhibits lung cancer cell growth in vitro and in vivo [2]. SLC6A4 is a serotonin reuptake transporter that recycles serotonin from the synaptic cleft. According to the results, SLC6A4 overexpression in non-small cell lung cancer is linked to a dismal prognosis because it turns on the cMyc oncogene [29]. High expression in the binding subgroup of CSF3 and granulocyte-colony stimulating factor 3 receptor (CSF3R) has been associated with a good prognosis in LUAD and a poor prognosis in LUSC. The highly expressed binding subset of CSF3 and CSF3R indicates an unfavorable prognosis, which is co-regulated by all common upregulated genes associated with inflammatory and signal transduction [30].

In our study, the role of genes that may be associated with SCLC was investigated using bioinformatic analyses taken from public databases. In SCLC, the expression and functions of genes related to immunity in immune cells that had moved into the tumor and to inflammation in fibroblasts were studied. Additional studies will be conducted using SCLC cohorts with larger patient groups, which should support the results obtained from this study. In order to confirm our findings, the expression levels of genes identified as significant by RNA sequencing should be investigated using real-time quantitative PCR as an alternative tool.

So that we can figure out whether the SLCL prognostic process is good or bad, we look at the pathways where the differentially expressed genes are enriched for both up-regulated and down-regulated genes separately. These genes are increased in SCLC, and their products are involved in cellular processes such as mitosis, the cell

cycle, DNA repair, transcriptional control, and the p53 signaling pathway. These pathways are the primary targets of anti-cancer medications and are well-known to have a significant role in cancer etiology. It has been discovered that the products of genes that are downregulated in SCLC are linked to the control of the immune response. Only 75 SCLC patients and 7 control samples were used in this study's experimental data set for gene expression analysis in SCLC lung tissue samples through Array Express data. This shows that the sample size for the control group is small. The Human Protein Atlas database was used to find out which cell types express these protein-coding genes. The Array Express platform and the GEO2R program through the GEO database were used to profile and analyze gene expression. These databases included in the study are web-accessible and offer easy accessibility and rapid evaluation of data for bioinformatics analyses.

In the future, we expect that by analyzing RNA sequencing and proteogenomic data for this fatal illness, more effective treatment regimens based on the molecular processes of SCLC will be developed. Also, genomic and transcriptomic data from gene expression profiling show that they can be used as molecular biomarkers for targeted therapies in SCLC.

V. CONCLUSION

It is well established that the tumor microenvironment is a crucial factor in tumor development and progression. We determined that the tumor microenvironment consists of stromal cells (fibroblasts and mesenchymal cells), macrophages, neutrophils, and Type I and Type II alveolar cells, in addition to cancer cells. Based on the data we collected, new therapeutic approaches can be made using potential prognostic and predictive biomarkers that target all of these cells in the tumor microenvironment to find out how SCLC grows.

Ethics Committee Approval: Since this study was obtained from bioinformatics analysis data, ethics committee approval is not required.

REFERENCES

[1] Liang, J., Guan, X., Bao, G., Yao, Y., & Zhong, X. (2022). Molecular subtyping of small cell lung cancer. *Semin Cancer Biol*, 86(Pt 2), 450-462.

[2] Li, C., Lei, S., Ding, L., Xu, Y., Wu, X., Wang, H., Zhang, Z., Gao, T., Zhang, Y., Li, L. (2023). Global burden and trends of lung cancer incidence and mortality. *Chin Med J*, 136(13), 1583-1590.

[3] Thandra, K.C., Barsouk, A., Saginala, K., Aluru, J.S., Barsouk, A. (2021). Epidemiology of lung cancer. *Contemp Oncol,* 25(1), 45-52.

[4] Siegel, R.L., Miller, K.D., Wagle, N.S., Jemal, A. (2023). Cancer statistics, 2023. *CA Cancer J Clin*, 73(1), 17-48.

[5] Rudin, C. M., Brambilla, E., Faivre-Finn, C., & Sage, J. (2021). Small-cell lung cancer. *Nat Rev Dis Primers,* 7(1), 3.

[6] Abe, Y., Tanaka, N. (2016). The Hedgehog Signaling Networks in Lung Cancer: The Mechanisms and Roles in Tumor Progression and Implications for Cancer Therapy. *Biomed Res Int,* 2016(7969286), 1-11.

[7] van Meerbeeck, J. P., Fennell, D. A., & De Ruysscher, D. K. (2011). Small-cell lung cancer. *Lancet,* 378(9804), 1741-1755.

[8] Gupta, S., Kass, G.E.N., Szegezdi, E., Joseph, B. (2009). The mitochondrial death pathway: a promising therapeutic target in diseases. *J Cell Mol Med,* 13(6), 1004-33.

[9] Li, Q., Wang, R., Yang, Z., et al. (2022). Molecular profiling of human non-small cell lung cancer by single-cell RNA-seq. *Genome Med,* 14(1), 87.

[10] Tian, Y., Li, Q., Yang, Z., et al. (2022). Single-cell transcriptomic profiling reveals the tumor heterogeneity of small-cell lung cancer. *Signal Transduct Target Ther*, 7(1), 346.

[11] Meijer, J. J., Leonetti, A., Airo, G., et al. (2022). Small cell lung cancer: Novel treatments beyond immunotherapy. *Semin Cancer Biol*, 86(Pt 2), 376-385.

[12] Wang, Y., Zou, S., Zhao, Z., Liu, P., Ke, C., & Xu, S. (2020). New insights into small-cell lung cancer development and therapy. *Cell Biol Int,* 44(8), 1564-1576.

[13] Agapito, G., Milano, M., Cannataro, M. (2022). A statistical network pre-processing method to improve relevance and significance of gene lists in microarray gene expression studies. *BMC Bioinformatics,* 23(6):393.

[14] Hayashi, R., & Inomata, M. (2022). Small cell lung cancer; recent advances of its biology and therapeutic perspective. *Respir Investig*, 60(2), 197-204.

[15] Yuan, M., Zhao, Y., Arkenau, H. T., Lao, T., Chu, L., & Xu, Q. (2022). Signal pathways and precision therapy of small-cell lung cancer. *Signal Transduct Target Ther*, 7(1), 187.

[16] Li, S., Zhang, J., Zhao, Y., Wang, F., Chen, Y., & Fei, X. (2018). miR-224 enhances invasion and metastasis by targeting HOXD10 in non-small cell lung cancer cells. *Oncol Lett*, 15(5), 7069-7075.

[17] Liu, H., Li, T., Ye, X., & Lyu, J. (2021). Identification of Key Biomarkers and Pathways in Small-Cell Lung Cancer Using Biological Analysis. *Biomed Res Int*, 2021, 5953386.

[18] Yu, J., Hou, M., & Pei, T. (2020). FAM83A Is a Prognosis Signature and Potential Oncogene of Lung Adenocarcinoma. *DNA Cell Biol*, 39(5), 890-899.
[19] Bai, S., Zhao, H., Zeng, X., et al. (2021). FAM83A-AS1 Promotes Tumor Progression Through MET Signaling in Lung Adenocarcinoma. *Research Square.* 1-19
[20] Cui, F., Zhou, Q., Xiao, K., & Ma, S. (2020). The MicroRNA hsa-let-7g Promotes Proliferation and Inhibits Apoptosis in Lung Cancer by Targeting HOXB1. *Yonsei Med J*, 61(3), 210-217.
[21] Gao, L., Hu, Y., Tian, Y., et al. (2019). Lung cancer deficient in the tumor suppressor GATA4 is sensitive to TGFBR1 inhibition. *Nat Commun,* 10(1), 1665.
[22] Yang, D., Liu, M., Jiang, J., et al. (2022). Comprehensive Analysis of DMRT3 as a Potential Biomarker Associated with the Immune Infiltration in a Pan-Cancer Analysis and Validation in Lung Adenocarcinoma. *Cancers (Basel)*, 14(24).
[23] Zhang, S., Li, M., Ji, H., & Fang, Z. (2018). Landscape of transcriptional deregulation in lung cancer. *BMC Genomics*, 19(1), 435.
[24] Zhang, W., Wu, C., Zhou, K., et al. (2022). Clinical and immunological characteristics of TGM3 in pan-cancer: A potential prognostic biomarker. *Front Genet,* 13, 993438.
[25] Xu, L., Qin, Y., Sun, B., et al. (2020). Involvement of CHP2 in the Development of Non-Small Cell Lung Cancer and Patients' Poor Prognosis. *Appl Immunohistochem Mol Morphol*, 28(9), 678-686.
[26] Takakura, S., Kohno, T., Manda, R., Okamoto, A., Tanaka, T., & Yokota, J. (2001). Genetic alterations and expression of the protein phosphatase 1 genes in human cancers. *Int J Oncol*, 18(4), 817-824.
[27] Xia, D., Chen, Z., & Liu, Q. (2021). Circ-PGC increases the expression of FOXR2 by targeting miR-532-3p to promote the development of non-small cell lung cancer. *Cell Cycle*, 20(21), 2195-2209.
[28] Li, B., Meng, Y. Q., Li, Z., et al. (2019). MiR-629-3p-induced downregulation of SFTPC promotes cell proliferation and predicts poor survival in lung adenocarcinoma. *Artif Cells Nanomed Biotechnol*, 47(1), 3286-3296.
[29] Pappula, A. L., Gibson, L. N., Bouley, R. A., & Petreaca, R. C. (2022). In silico analysis of a SLC6A4 G100V mutation in lung cancers. *MicroPubl Biol,* 2022.
[30] Huang, X., Hu, P., & Zhang, J. (2020). Genomic analysis of the prognostic value of colony-stimulating factors (CSFs) and colony-stimulating factor receptors (CSFRs) across 24 solid cancer types. *Ann Transl Med*, 8(16), 994.