Asian J Biol Life Sci., 2025; 14(2):400-409.
https://ajbls.com

Research Article

Identification of Hub Genes and Therapeutic Targets in Lung Cancer Using Software Tools

S. Anbuselvi*, S. Saranya

Department of Biotechnology, Bharath Institute of Higher Education and Research, Chennai, Tamil Nadu, INDIA.

ABSTRACT

Background: Lung cancer remains one of the leading causes of cancer-related mortality globally, necessitating a deeper understanding of its molecular underpinnings. **Objectives:** This study aimed to identify hub genes associated with lung cancer through a systematic analysis of existing literature and databases. **Materials and Methods:** We curated a comprehensive Gene Expression Omnibus (GEO) dataset and identified key hub genes linked to lung cancer via PubMed resources. To assess the relevance of these hub genes, we conducted Protein-Protein Interaction (PPI) analysis through the DAVID database, selecting those with high enrichment values. Functional enrichment analysis was performed using DAVID, SHINY GO, and GO NET DICE to elucidate the biological processes and pathways associated with the identified hub genes. Additionally, we employed ChEMBL, Pharos, and Broad tools to assess druggability, integrating chemical, bioactivity, and genomic data. Functional gene partners were grouped to provide a clearer understanding of the interaction networks. **Results:** The genes were then ranked based on their involvement in various molecular functions, yielding insights into their potential roles in the pathology of lung cancer. **Conclusion:** This comprehensive analysis underscores critical gene interactions and functional pathways, offering promising targets for future research and therapeutic intervention in the treatment of lung cancer.

Keywords: Lung Cancer, Gene, Gene Expression Omnibus, Pathology, Therapeutic.

Corresponding Author:
Dr. S. Anbuselvi
Department of Biotechnology, Bharath Institute of Higher Education and Research, Chennai-73, Tamil Nadu, INDIA.
Email: drsanbuselvi@gmail.com

Received: 12-02-2025;
Revised: 09-04-2025;
Accepted: 27-06-2025.

INTRODUCTION

Lung cancer remains the leading cause of cancer-related deaths worldwide, presenting significant challenges in both clinical management and research. Non-Small Cell Lung Cancer (NSCLC), which accounts for over 80% of lung cancer cases, exemplifies the complexity of this disease through its remarkable genetic and phenotypic diversity. This heterogeneity, driven by both intertumor and intratumor variations, complicates diagnosis, treatment planning, and prognosis. While advances in genomic technologies-particularly Single-Cell RNA sequencing (scRNA-seq)-have revolutionized our understanding of tumor biology, they also highlight the ongoing challenges of precision medicine in the context of NSCLC. These insights offer opportunities to improve treatment outcomes, but also underline the need for continued research and innovation in therapeutic strategies.[1]

NSCLC is characterized by considerable heterogeneity at both the intertumor and intratumor levels. Intertumor heterogeneity refers to the differences observed between tumors of the same histological type in different patients, while intratumor heterogeneity involves the genetic and phenotypic variability within a single tumor. These variations are driven by the genetic evolution of the tumor, which is influenced by factors such as mutation accumulation, environmental exposures (e.g., smoking), and therapeutic pressures. Intratumor heterogeneity is particularly concerning, as it means that within one tumor, different subclones of cancer cells may harbor distinct mutations, leading to differential responses to treatments. For instance, subclones with specific mutations might be resistant to chemotherapy or targeted therapies, while others remain sensitive. These diverse subpopulations of cancer cells complicate treatment planning, as tumors may evolve under treatment pressure, leading to drug resistance or disease progression. Personalized medicine, in which treatment is tailored to the specific genetic profile of an individual's tumor, is crucial for overcoming these challenges. To further complicate matters, tumor evolution is not static, and tumors often acquire new mutations over time. This necessitates continuous monitoring of tumor evolution and frequent adjustments in treatment regimens to maintain therapeutic efficacy.[2]

The Tumor Microenvironment (TME) plays a crucial role in tumor progression, metastasis, and response to treatment. TME is a complex, dynamic environment consisting of a variety of cellular

OPEN ACCESS

ScienScript

DOI: 10.5530/ajbls.20251509

Publishing Partner : ScienScript Digital. [www.scienscript.com.sg]

and extracellular components, including fibroblasts, immune cells, endothelial cells, and Extracellular Matrix (ECM) proteins. These components interact with tumor cells, influencing their behavior, survival, and ability to resist treatment. One of the most important cellular players in the TME is the Cancer-Associated Fibroblast (CAF), which actively contributes to ECM remodeling, tumor growth, and metastasis. CAFs also facilitate angiogenesis, thereby supporting tumor vascularization and enabling tumor cells to access nutrients and oxygen. Moreover, fibroblasts and stromal cells interact with cancer cells to create a tumor niche that promotes cancer cell proliferation, survival, and invasion, making CAFs an important target for therapeutic strategies. Macrophages are often polarized into two distinct types: pro-inflammatory M1 macrophages, which can inhibit tumor growth, and anti-inflammatory M2 macrophages, which support tumor progression. The balance between these different macrophage subtypes plays a significant role in regulating tumor progression. Similarly, the role of neutrophils in the TME is complex-they can either inhibit or promote tumor progression depending on their activation state, complicating their role in cancer therapy. Understanding these complex interactions within the TME is essential for developing strategies that can manipulate the TME to enhance the efficacy of treatment.

Technological advancements, such as single-cell RNA sequencing (scRNA-seq), have greatly enhanced our understanding of the cellular composition of tumors and the TME. scRNA-seq allows for the detailed examination of gene expression profiles at the individual cell level, revealing previously unrecognized cellular heterogeneity within both tumor cells and stromal components. Through this technology, researchers can identify distinct tumor subpopulations, map cellular interactions, and uncover Differentially Expressed Genes (DEGs) that might serve as potential biomarkers for diagnosis or therapeutic targets. For example, scRNA-seq has been used to identify novel subclones within tumors that were previously undetectable through bulk sequencing techniques, enabling a more comprehensive view of tumor evolution and resistance mechanisms. This technology is paving the way for personalized treatment strategies, as it enables clinicians to design more targeted therapies based on the unique molecular profile of a patient's tumor. Additionally, scRNA-seq offers the potential for early detection of therapeutic resistance and real-time monitoring of tumor dynamics, helping clinicians to adapt treatment strategies as tumors evolve.[3,4]

Despite significant advances in our understanding of NSCLC, there remain several critical challenges in its treatment. One of the major hurdles is variable responses to immunotherapy and the absence of reliable biomarkers to predict which patients will benefit from such treatments. For instance, while Immune Checkpoint Inhibitors (ICIs) like PD-1/PD-L1 blockers have shown promise, only a subset of patients responds to these therapies, and many develop resistance over time. Identifying biomarkers to predict patient response to ICIs would significantly improve treatment outcomes and avoid unnecessary side effects for non-responders. Furthermore, treatment resistance, especially to targeted therapies, remains a persistent issue. As tumor cells acquire new mutations, they may bypass the effects of targeted therapies, requiring novel strategies to overcome or prevent resistance.

Another emerging area of research is the role of microbiomes in cancer progression and treatment response. Recent studies suggest that the microbiota-consisting of diverse bacterial populations in the body-can influence tumor behavior and the effectiveness of treatments. Microbial composition may modulate immune responses or affect the metabolism of drugs, suggesting that the microbiome could be leveraged to improve treatment efficacy. Understanding these interactions between the microbiome and the tumor could lead to novel strategies for optimizing cancer therapy. The complexity of NSCLC, driven by both tumor heterogeneity and the dynamic TME, poses significant challenges for treatment. Advances in technologies like scRNA-seq have enhanced our understanding of tumor biology, offering new insights into cellular diversity and tumor evolution. These tools enable more personalized approaches to treatment but also highlight the need for continued research into mechanisms of resistance, predictive biomarkers, and the impact of the microbiome. As our understanding of the molecular and cellular landscape of NSCLC deepens, it holds the promise of improving patient outcomes and informing the development of novel therapeutic strategies. A comprehensive approach that integrates genomic technologies, a better understanding of the TME, and personalized medicine will be key to addressing the complex challenges of NSCLC.[5] This study aims to provide a foundation for developing potential diagnostic markers and therapeutic targets for lung cancer.

MATERIALS AND METHODS

The immune lung cancer related 20 HUB genes were selected from Gene Expression Omnibus (GEO) datasets and their descriptions. These genes were experimentally proven to be involved in the pathogenesis of immune lung cancer and extensively analysed using bioinformatics tools, gene set enrichment analysis, etc., Identification of crucial HUB genes as effective protein targets from this list of HUB genes are further analysed using bioinformatics approaches. Our approach involved ranking of HUB genes through PPI analysis and functional enrichment across different GO databases and identifying crucial HUB genes through number of functional involvements.

PPI analysis using Stringbase

Stringbase version 11.5 is an open-source web based database available through www.string-db.org. It depicts functional association between two proteins jointly contributing to a

specific function with PPI interaction score values. It predicts protein-protein interactions scores through computational analysis from interactions aggregated from other (primary) databases and experimental results of genomic content predictions, high-throughput lab experiments, conserved co-expressions, automated text mining etc. In string base, the edges are representation between two interacting nodal proteins as predicted edges derived based on experimental evidence and protein associations from 12 different biological data sources. 20 hub genes associated with lung cancer were used as keywords and node to node interaction data were saved into excel sheet by downloading tabular text outputs.[6]

Functional Enrichment Analysis using DAVID

The Database for Annotation, Visualization and Integrated Discovery (DAVID) is available as open source through https://david.ncifcrf.gov/home.jsp. It is developed as a functional annotation tool for using a large set of genes. It retrieves data by listing functional enriched related genes from BioCarta and KEGG pathway maps, linking gene-disease associations, highlighting functional domains and motifs etc., 20 hub genes associated with lung cancer were used as gene lists in DAVID search by converting the gene list using DAVID gene ID converter tool and data downloaded in MS-Excel for molecular functional clustering and enrichment analysis.[7]

Functional enrichment analysis and grouping of functional gene partners using ShinyGO

ShinyGO version 0.76 is an open source web based GO tool used for high level functional enrichment available through http://bioinformatics.sdstate.edu/go/ which produces protein interaction network visualizations, clustering trees, pathways etc. based on annotation from Ensembl. The number of folds in functional enrichments with number of genes involved with False Discovery Rate (FDR) values can be graphically represented. Further, the categorization of numbers and groups of genes involved in high level molecular functional enrichment are represented in tabular format downloadable directly from the database. The correlation of significant functional pathway enrichment is represented through hierarchical tree clustering and results were downloaded in PNG format from database.[8]

Functional enrichment analysis and grouping of functional gene partners using GoNET DICE

The DICE (Database of Immune Cell Expression) is freely available web source at http://tools.dice-database.org/GOnet/. It can perform analysis of GO term annotation or gene enrichment analysis by taking a list or set of human genes. It can depict functionally interconnected user-friendly network of submitted set of gene list.[9]

Ranking of HUB genes and functional enrichment analysis

The data from DAVID, ShinyGo and GONET were comparable based on the involvement of the genes in various functions. Based on the commonality of functions, the different genes were grouped, and ranking reveals the influence of a single protein in various functions.

ChEMBL

ChEMBL is an open-source database https://www.ebi.ac.uk/chembl/ that serves as a manually curated resource of bioactive molecules, particularly focusing on drug-like small molecules. It contains 2-D structures, calculated properties (such as logP, molecular weight, and Lipinski parameters), and abstracted bioactivity data (including binding constants, pharmacology, and ADMET profiles). This search yields over 10 targets related to the entered gene. By clicking on the full form of the gene, users can access detailed information, including ligand efficiencies, approved drugs, and clinical candidates associated with those targets. This functionality enhances the ability to identify potential therapeutic avenues for lung cancer based on molecular interactions and drug development status.[10]

PHAROS

Pharos is the user interface for the Knowledge Management Center (KMC), part of theIlluminating the Druggable Genome (IDG) program funded by the NIH Common Fund (Grant No.**1U24CA224370-01**). The KMC aims to create a comprehensive knowledge base for the Druggable Genome (DG), focusing on poorly characterized or unannotated regions. It specifically emphasizes three commonly drug-targeted protein classes: G-Protein-Coupled Receptors (GPCRs), Ion Channels (ICs), and kinases. This feature enhances the ability to identify potential therapeutic targets and their associated drug development status for lung cancer.[11]

BROAD

CLUE is a cloud-based software platform designed for the analysis of perturbational datasets generated from gene expression (L1000) and proteomic (P100 and GCP) assays. The current DATA VERSION: Beta and theSOFTWARE VERSION: 1.1.1.43. Recent technological advancements have dramatically increased the availability of high-dimensional perturbational datasets to the biomedical community. This user-friendly interface facilitates easy access to relevant datasets, empowering researchers to derive insights and formulate hypotheses related to lung cancer.

RESULTS AND DISCUSSION

The gene clustering analysis of 147 HUB genes in DAVID were used in identifying functional enrichments. We found 107 genes in DAVID database and found 37 genes involved in various

Table 1: Functional enrichment analysis of 147 lung cancer genes in DAVID. A total of 9 gene clusters under different Gene Ontology terms were shown with functional fold enrichment.

Sl. No.	Term	Genes	Fold Enrichment	FDR
1	TOPO_DOM:Extracellular	MRC1, CD84, CXCR4, TIGIT, LAG3, IL2RA, CCR7, CXCR2, CD74, CD79A, CXCR1, CD19, CD79B, SELL, CD79B, CD8A, CD3D, CD3E, CD3G	6.530932	2.27E-12
2	TOPO_DOM:Cytoplasmic	MRC1, CD84, CXCR4, TIGIT, LAG3, IL2RA, CCR7, CXCR2, CD74, CD79A, CXCR1, CD19, CD79B, SELL, CD79B, CD8A, CD3D, CD3E, CD3G	5.009113	1.29E-10
3	KW-1003~Cell membrane	MRC1, CD84, CXCR4, TIGIT, LAG3, CCR7, CXCR2, CD74, CD79A, CXCR1, CD19, CD79B, SELL, CD79B, CD8A, CD3D, CD3E, CD3G	3.988096	1.06E-08
4	KW-1015~Disulfide bond	MRC1, CD84, CXCR4, VCAN, TIGIT, LAG3, IL2RA, CCR7, CXCR2, CD74, CD79A, CXCR1, CD19, CD79B, SELL, CD79B, CD8A, CD3D, CD3E, CD3G	3.648746	2.02E-10
5	TRANSMEM:Helical	MRC1, CD84, CXCR4, TIGIT, LAG3, IL2RA, CCR7, CXCR2, CD74, CD79A, CXCR1, CD19, CD79B, SELL, CD79B, CD8A, CD3D, CD3E, CD3G	3.588055	1.93CD79A577544735E-08
6	GO:0005886~plasma membrane	MRC1, CD84, CXCR4, TIGIT, LAG3, IL2RA, CCR7, CXCR2, CD74, CD79A, CXCR1, CD19, CD79B, SELL, CD79B, CD8A, CD3D, CD3E, CD3G	3.548399	1.95E-08
7	KW-1133~Transmembrane helix	MRC1, CD84, CXCR4, TIGIT, LAG3, IL2RA, CCR7, CXCR2, CD74, CD79A, CXCR1, CD19, CD79B, SELL, CD79B, CD8A, CD3D, CD3E, CD3G	2.357557	1.80E-06
8	KW-0812~Transmembrane	MRC1, CD84, CXCR4, TIGIT, LAG3, IL2RA, CCR7, CXCR2, CD74, CD79A, CXCR1, CD19, CD79B, SELL, CD79B, CD8A, CD3D, CD3E, CD3G	2.334144	1.80E-06
9	KW-0472~Membrane	MRC1, CD84, CXCR4, TIGIT, LAG3, IL2RA, CCR7, CXCR2, CD74, CD79A, CXCR1, CD19, CD79B, SELL, CD79B, CD8A, CD3D, CD3E, CD3G	2.043753	5.41E-05

functions, out of which we selected top 20 genes for further studies. We found 8 clusters and 107 GO terms in DAVID and selected 20 HUB genes and found 9 functionally enriched GO terms with different gene clusters associated with Extracellular, cell membrane, disulfide bond and cytoplasmic receptor binding activities (Table 1). String base exhibited 118 multiple edges with 20 multiple nodes to node interactions in which most of the GO terms in network clusters were found associated with C-C Chemokine receptor activity and binding, and T cell receptor binding as shown in Figure 1. Local network clustering found 20 HUB genes described under 15 different GO terms were found, in which maximum nodal interaction with highest number of functional enrichments were associated with Chemokine receptor binding. Each different terms contained a group of 2-16 genes involving in Signaling receptor activity, MHC protein binding, Protein binding and T cell receptor binding are shown in Table 2.

These 9 different categories of GO terms were reportedly enriched with 5-9 genes in each group contributing functional descriptions and were reported through KEGG pathways whose details are shown in Table 2. The signalling pathways in SHINY GO contains the functional network relationships of top ranked 20 genes were described under 20 GO terms involving various protein binding and metabolic receptors in Figure 2.

We tested functional enrichment of 20 HUB genes in Shiny GO, which were shown in Table 3. The functional pathways of top ranked 20 HUB genes in SHINY GO 6 different group of genes (each group 1-6 genes) with a maximum of 200% functional enrichments in C-C Chemokine receptor activity, G protein coupled chemo attractant receptor activity. carbohydrate metabolisms, T cell receptor binding and C-C Chemokine binding. The process of C-C Chemokine receptor activity exhibited highest functional fold enrichments (>200%) genes were found. 1-6 gene clusters were found with 50-200-fold functional

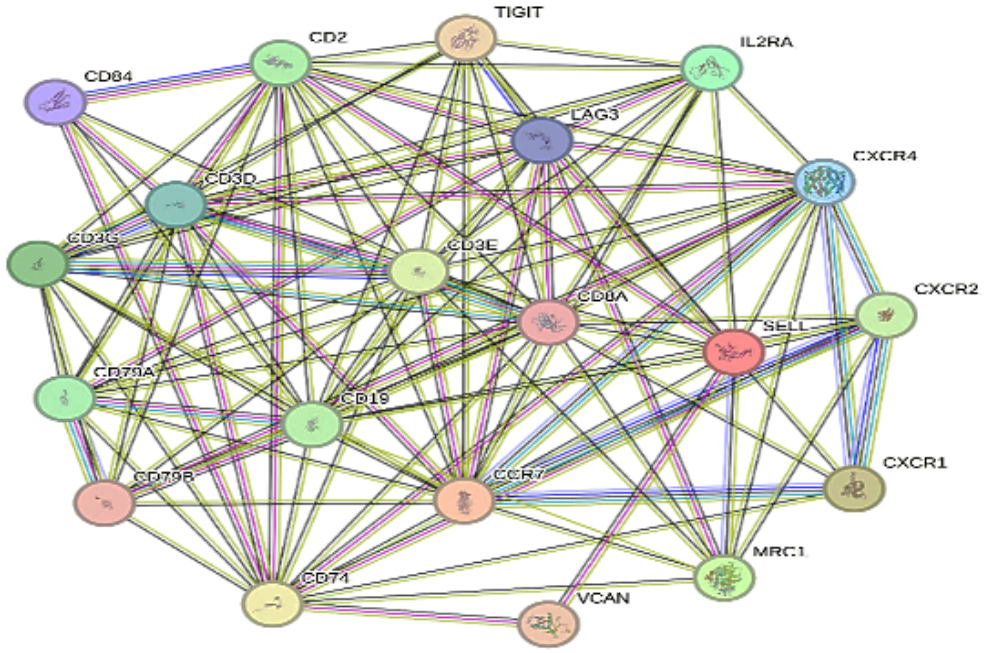

Figure 1: PPI interactions of top 20 HUB Genes interacting proteins in stringbase. Number of Nodes : 20, Number of Edges : 118, Average Node Degree:11.8, Avg. Local Clustering Coefficient : 0.777, Expected Number of Edges : 12, PPI Enrichment *p*-value : < 1.0e-16.

Table 2: Top ranked 20 HUB genes and their related GO terms in KEGG network clusters in String base.

Sl. No.	Description	GENES
1	Signaling receptor activity	CD74,LAG3,CD79A,CCR7,CXCR1,CXCR2,CD3E,CD2,IL2RA,CD79B,CD8A,CXCR4,CD3G,MRC1
2	Transmembrane signaling receptor activity	CD74,LAG3,CD79A,CCR7,CXCR1,CXCR2,CD3E,IL2RA,CD79B,CXCR4,CD3G,MRC1
3	Cytokine receptor activity	CD74,CCR7,CXCR1,CXCR2,IL2RA,CXCR4
4	Cytokine binding	CD74,CCR7,CXCR1,CXCR2,IL2RA,CXCR4
5	C-C chemokine receptor activity	CCR7,CXCR1,CXCR2,CXCR4
6	C-C chemokine binding	CCR7,CXCR1,CXCR2,CXCR4
7	C-X-C chemokine receptor activity	CXCR1,CXCR2,CXCR4
8	interleukin-8 receptor activity	CXCR1,CXCR2
9	MHC protein binding	CD74,LAG3,CD8A
10	interleukin-8 binding	CXCR1,CXCR2
11	Protein binding	CD74,LAG3,CD79A,SELL,CCR7,CXCR1,CD84,CXCR2,CD3E,CD2,IL2RA,CD79B,CD8A,CXCR4,TIGIT,CD3G
12	MHC class II protein binding	CD74,LAG3
13	Identical protein binding	CD74,CD79A,CD84,CD3E,CD2,CD79B,CXCR4,TIGIT,CD3G
14	T cell receptor binding	CD3E,CD3G
15	MHC protein complex binding	CD74,CD8A

Table 3: Functional Enrichment Analysis of the top 20 HUB genes interacting proteins in Shiny GO.

Sl. No.	Pathway	Genes	Fold Enrichment	Enrichment FDR
1	GO:0016493 C-C chemokine receptor activity	CXCR4 CCR7 CXCR1 CXCR2	198.9652	5.22E-08
2	GO:0019957 C-C chemokine binding	CXCR4 CCR7 CXCR1 CXCR2	190.675	5.22E-08
3	GO:0001637 G protein-coupled chemoattractant receptor activity	CXCR4 CCR7 CXCR1 CXCR2	176.0077	5.71E-08
4	GO:0004950 chemokine receptor activity	CXCR4 CCR7 CXCR1 CXCR2	176.0077	5.71E-08
5	GO:0019956 chemokine binding	CXCR4 CCR7 CXCR1 CXCR2	138.6727	1.40E-07
6	GO:0042608 T cell receptor binding	CD3G CD3E	99.48261	0.000745
7	GO:0050998 nitric-oxide synthase binding	CD74	88.00385	0.025754
8	GO:0070492 oligosaccharide binding	SELL	76.27	0.028889
9	GO:0004896 cytokine receptor activity	CD74 CXCR4 CCR7 CXCR1 CXCR2 IL2RA	67.29706	5.11E-09
10	GO:0005537 mannose binding	MRC1	49.7413	0.041886
11	GO:0030021 extracellular matrix structural constituent conferring compression re	VCAN	49.7413	0.041886
12	GO:0030159 signaling receptor complex adaptor activity	CD3G CD3E	44.86471	0.003343
13	GO:0015026 coreceptor activity	CD8A CXCR4	44.00192	0.003343
14	GO:0019955 cytokine binding	CD74 CXCR4 CCR7 IL2RA CXCR1 CXCR2	42.11227	5.22E-08
15	GO:0046625 sphingolipid binding	SELL	39.45	0.048919
16	GO:0051861 glycolipid binding	SELL	39.45	0.048919
17	GO:0005540 hyaluronic acid binding	VCAN	38.135	0.049408
18	GO:0042287 MHC protein binding	CD74 LAG3 CD8A	34.66818	0.000366
19	GO:0008528 G protein-coupled peptide receptor activity	CXCR4 CCR7 CXCR1 CXCR2	29.33462	5.24E-05
20	GO:0001653 peptide receptor activity	CXCR4 CCR7 CXCR1 CXCR2	28.24815	5.24E-05

enrichment in C-C Chemokine receptor binding, T Cell receptor binding, Nitric oxide synthase binding, Oligosaccharide binding, Mannose binding and Sphingolipid binding. More than 10 HUB were reported with 50-fold functional enrichment in pathways of cancers. In Shiny GO hierarchical clustering tree summarizing the 15 GO Term clusters in which bigger dots indicate more significant P value were shown in Figure 2.

In GO Net dice, out of 20 HUB genes we found 17 HUB genes which are involved in transmebrane signaling receptor activity, cytokine receptor activity and signalling receptor activity Figure 3. Based on PPI analysis and functional enrichment results across these GO databases, we ranked these genes and shortlisted 20 HUB genes based on their total number of functional involvements and their related functions with references shown in Table 4. Out of these 20 HUB genes 15 HUB genes are approved for drugs in which CXCR4 gene is approved in all three databases (chEMBL, Pharos and Broad) in both Active ligands/Comounds and Clinical trial phase. CXCR4, CD2, CD74, LAG3, IL2RA and SELL genes have drugable property to cure lung cancer disease as shown in Table 5.

Bioinformatics tools and web-based Gene Ontology (GO) databases have revolutionized the identification and validation of potential therapeutic targets for complex diseases such as Immune Lung Cancer (ILC). In our study, the integration of GO databases and Protein-Protein Interaction (PPI) networks was pivotal in elucidating the functional roles of immune-related genes and identifying key therapeutic targets. Functional network analysis, utilizing platforms like ShinyGO, DAVID, and GoNet, allowed us to systematically categorize and rank genes based on their involvement in various biological processes and pathways. This

Table 4: **Ranking of the top-ranked HUB Genes based on their involvement in the total number of functions.**

Sl. No.	Genes	DAVID Rank	Genes	GO NET Rank	Genes	SHINY GO Rank
1	CXCR4	9	CXCR2	15	CXCR4	10
2	LAG3	9	CXCR1	15	CXCR1	9
3	CD84	9	CXCR4	13	CXCR2	9
4	CD74	9	CD74	12	CCR7	9
5	CD3G	9	CCR7	12	CD74	4
6	CD3E	9	CD3G	6	SELL	3
7	CD3D	9	CD3E	6	CD3G	2
8	CD2	9	LAG3	5	CD3E	2
9	CD79B	9	IL2RA	5	IL2RA	2
10	CD79A	9	CD79B	5	CD8A	2
11	CXCR1	9	CD79A	5	VCAN	2
12	SELL	9	CD3D	5	LAG3	1
13	CD8A	9	CD2	4	MRC1	1
14	CD19	9	MRC1	3	CD3D	
15	CXCR2	9	CD8A	3	CD79B	
16	CCR7	9	TIGIT	1	CD79A	
17	MRC1	9	CD84	1	CD2	
18	TIGIT	9	VCAN		CD84	
19	IL2RA	8	SELL		TIGIT	
20	VCAN	1	CD19		CD19	

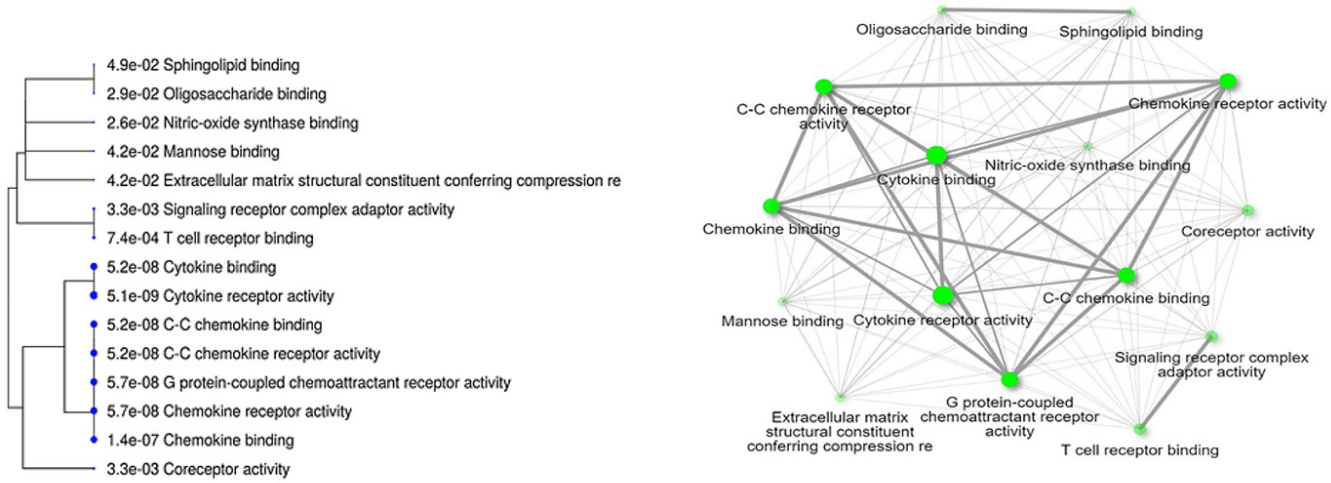

Figure 2: (a) hierarchical clustering tree summarizing the correlation among significant pathways of top 20 HUB genes interaction in ShinyGO. Pathways with many shared genes were clustered together. Bigger dots indicate more significant P-values. (b): The signaling pathways and their networks of functional relationships among the 20 top ranked genes in shiny go.

Table 5: List of 20 HUB gene interacting drugs are analysis in chEMBL, Pharos and Broad databases to cure the immune lung cancer disease.

Sl. No.	Gene Name	Functions of Hub Genes	chEMBL		Pharos		Broad	
			Active ligands	Clinical trail phase	Active ligands	Clinical trail phase	Active ligands	Clinical trail phase
1	CXCR4	Involved in chemokine receptor	YES	YES	YES	YES	YES	YES
2	CD2	Involved in T cell surface antigen	YES	YES	YES	YES	NIL	NIL
3	CD74	Involved in HLA class II histocompatibility antigen	YES	YES	YES	NIL	NIL	NIL
4	LAG3	Involved in lymphocyte activation	YES	YES	NIL	YES	NIL	NIL
5	IL2RA	Involved in interleukin 2 receptor	YES	YES	NIL	YES	NIL	NIL
6	SELL	Involved in L selectin	NIL	YES	YES	NIL	YES	NIL
7	CD3E	Involved in T cell surface glycoprotein	NIL	YES	NIL	YES	NIL	NIL
8	CD3G	Involved in T cell surface glycoprotein	NIL	YES	NIL	YES	NIL	NIL
9	CD19	Involved in B lymphocyte antigen	NIL	YES	NIL	YES	NIL	NIL
10	CXCR2	Involved in chemokine receptor	NIL	NIL	YES	NIL	YES	NIL
11	CD79B	Involved in B cell antigen receptor	NIL	YES	NIL	YES	NIL	NIL
12	CCR7	Involved in chemokine receptor	YES	NIL	YES	NIL	NIL	NIL
13	CXCR1	Involved in chemokine receptor	NIL	NIL	YES	YES	NIL	NIL
14	CD3D	Involved in T cell surface glycoprotein	NIL	NIL	NIL	YES	NIL	NIL
15	TIGIT	Involved in T cell immunoreceptor	NIL	YES	NIL	NIL	NIL	NIL
16	MRC1	Involved in macrophage mannose receptor	NIL	NIL	NIL	NIL	NIL	NIL
17	CD8A	Involved in T cell glycoprotein	NIL	NIL	NIL	NIL	NIL	NIL
18	VCAN	Involved in versican core protein	NIL	NIL	NIL	NIL	NIL	NIL
19	CD84	Involved in SLAM family member	NIL	NIL	NIL	NIL	NIL	NIL
20	CD79A	Involved in B cell antigen receptor	NIL	NIL	NIL	NIL	NIL	NIL

Figure 3: GO NET was used with top ranked 20 HUB genes and interactive network graph was derived. Functional network for 3 genes VCAN, CD19 and SELL were not described in GO NET DICE.

approach provided a comprehensive overview of gene function and interaction, leading to the identification of a top-ranked group of 20 genes with significant roles in ILC. By leveraging these bioinformatics resources, we were able to identify critical hub genes that are integral to the disease process and could serve as potential targets for novel therapeutic interventions.[12]

The chemokine receptors CXCR4, CXCR2, and CCR7 emerged as pivotal players in the context of ILC. These receptors are integral to various processes within the tumor microenvironment, including immune cell trafficking, angiogenesis, and metastasis. This receptor is well-documented for its role in cancer progression, particularly in lung and breast cancers. CXCR4 mediates the recruitment of cancer cells to metastatic sites and is associated with poor prognosis. Our findings support the CXCR2 is implicated in enhancing tumor growth and facilitating metastasis. Targeting CXCR2 could potentially disrupt the recruitment of immune cells to the tumor microenvironment, thereby impeding cancer progression and improving patient outcomes. CCR7 is crucial for lymphocyte trafficking and the formation of secondary lymphoid organs. Its involvement in tumor-related immune responses suggests that targeting CCR7 could influence the localization and activity of immune cells within the tumor microenvironment, potentially modulating the immune response and impacting tumor progression.[13]

Our study identified several immune-related hub genes with significant therapeutic implications. These genes play crucial roles in T-cell activation, cytokine signaling, and immune cell migration, making them potential targets for immunotherapy. CD2, CD3G, CD3E, and CD8A genes are integral components of the T-cell receptor complex, essential for T-cell activation and cytotoxic responses. Their upregulation in ILC underscores their involvement in modulating immune responses against tumors. Targeting these components could enhance T-cell activation and improve the efficacy of immunotherapeutic approaches.[14]

The findings from our study have significant clinical implications for the treatment of ILC. Identifying key genes involved in immune-related processes provides a foundation for developing targeted therapies and novel immunotherapeutic strategies. The potential of combining immune checkpoint inhibitors with targeted agents highlights the importance of personalized treatment strategies. By integrating biomarker data into clinical practice, we can optimize treatment regimens and improve patient outcomes. Future research should focus on validating the key genes identified in this study through experimental and clinical studies.

Our study highlights the critical role of functional network analysis and bioinformatics in identifying key genes involved in immune lung cancer. The identification of crucial genes, including CXCR4, CXCR2, CCR7, CD2, CD3G, CD3E, CD8A, SELL, LAG3, TIGIT, IL2RA, CD79A, and CD79B, provides a foundation for developing targeted therapies and novel immunotherapeutic strategies. Integrating bioinformatics approaches with clinical research is essential for advancing our understanding and treatment of complex diseases like immune lung cancer.[15-17]

The development of new immune checkpoint inhibitors and combination therapies holds promise for improving treatment outcomes and personalizing therapy.[18,19] By leveraging bioinformatics tools and web-based databases, we can uncover novel therapeutic targets and advance treatment strategies, paving the way for more effective and personalized therapies for immune lung cancer and beyond.[20,21] Future research will be crucial in validating these findings and exploring their full potential in clinical settings, ultimately leading to improved patient outcomes and enhanced therapeutic options.[22]

CONCLUSION

In conclusion, this study identifies key hub genes associated with lung cancer through a comprehensive integration of genomic data and literature. By analyzing datasets from the Gene Expression Omnibus, we elucidated critical protein-protein interactions and functional pathways relevant to the disease. We highlighted 20 top-ranked HUB genes, in which 15 HUB genes identified as promising candidates for diagnostic markers and therapeutic targets. This research not only enhances our understanding but also lays the groundwork for future investigations aimed at developing more effective treatment strategies.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest.

ABBREVIATIONS

NSCLC: Non-Small Cell Lung Cancer; **scRNA-seq:** single-cell RNA sequencing; **TME:** Tumor Microenvironment; **ECM:** Extracellular Matrix; **CAF:** Cancer-Associated Fibroblast; **DEGs:** Differentially Expressed Genes; **ICIs:** Immune Checkpoint Inhibitors; **PD-1:** Programmed Death-1; **PD-L1:** Programmed Death-Ligand 1; **GEO:** Gene Expression Omnibus; **PPI:** Protein-Protein Interaction; **DAVID:** Database for Annotation, Visualization and Integrated Discovery; **GO:** Gene Ontology; **KEGG:** Kyoto Encyclopedia of Genes and Genomes; **FDR:** False Discovery Rate; **DICE:** Database of Immune Cell Expression; **ChEMBL:** Chemical Database of the European Molecular Biology Laboratory; **IDG:** Illuminating the Druggable Genome; **KMC:** Knowledge Management Center; **NIH:** National Institutes of Health; **GPCRs:** G-Protein-Coupled Receptors; **ICs:** Ion Channels; **CLUE:** Connectivity Map; **L1000:** Library of Integrated Network-Based Cellular Signatures; **P100:**Phosphoproteomics; **GCP:** Gene Expression and Cellular Phenotype; **ILC:** Immune Lung Cancer; **MHC:** Major Histocompatibility Complex; **CXCR4:** C-X-C Motif Chemokine Receptor 4; **CXCR2:** C-X-C Motif Chemokine Receptor 2; **CXCR1:** C-X-C Motif Chemokine Receptor 1; **CCR7:** C-C Motif Chemokine Receptor 7; **CD2:** Cluster of Differentiation 2; **CD3G:** Cluster of Differentiation 3 Gamma; **CD3E:** Cluster of Differentiation 3 Epsilon; **CD3D:** Cluster of Differentiation 3 Delta; **CD8A:** Cluster of Differentiation 8 Alpha; **CD74:** Cluster of Differentiation 74; **CD79A:** Cluster of Differentiation 79A; **CD79B:**Cluster of Differentiation 79B; **CD19:** Cluster of Differentiation 19; **CD84:** Cluster of Differentiation 84; **LAG3:** Lymphocyte Activation Gene 3; **TIGIT:** T Cell Immunoreceptor with Ig and ITIM Domains; **IL2RA:** Interleukin 2 Receptor Alpha; **SELL:** Selectin L; MRC1:Mannose Receptor C-Type 1; **VCAN:** Versican; **HLA:** Human Leukocyte Antigen; **SLAM:** Signaling Lymphocytic Activation Molecule; **ADMET:** Absorption, Distribution, Metabolism, Excretion, and Toxicity; **logP:** Logarithm of the partition coefficient; **PNG:** Portable Network Graphics; **TIFF:** Tagged Image File Format; **dpi:** Dots per Inch; **MS-Excel:** Microsoft Excel.

REFERENCES

1. Becht E, Giraldo NA, Lacroix L, Buttard B, *et al.*, Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016; 17(1): 218. doi: 10.1186/s13059-016-1070-5. Erratum in: Genome Biol. 2016; 17(1): 249. doi: 10.1186/s13059-016-1113-y. PMID: 27765066; PMCID: PMC5073889.
2. Arigoni M, Ratto ML, Riccardo F, Balmas E, *et al.*, A single cell RNAseq benchmark experiment embedding "controlled" cancer heterogeneity. Sci Data. 2024; 11(1): 159. doi: 10.1038/s41597-024-03002-y. PMID: 38307867; PMCID: PMC10837414.
3. Maynard A, McCoach CE, Rotow JK, Harris L, *et al.*, Therapy-Induced Evolution of Human Lung Cancer Revealed by Single-Cell RNA Sequencing. Cell. 2020; 182(5): 1232-51.e22. doi: 10.1016/j.cell.2020.07.017. Epub 2020 Aug 20. PMID: 32822576; PMCID: PMC7484178
4. Zhao J, Zhu Q, Zhang Y, Li G, Zhang Y, Li F, Bian L. [Role of COX-2/PGE2/EP4 Axis-induced Macrophage Functional Activation
5. in NSCLC Development]. Zhongguo Fei Ai Za Zhi. 2024; 27(4): 245-56. Chinese. doi: 10.3779/j.issn.1009-3419.2024.101.05. PMID: 38769827; PMCID: PMC11110263
6. Der SD, Sykes J, Pintilie M, Zhu CQ, *et al.*, Validation of a histology-independent prognostic gene signature for early-stage, non-small-cell lung cancer including stage IA patients. J Thorac Oncol. 2014; 9(1): 59-64. doi: 10.1097/JTO.0000000000000042. PMID: 24305008.
7. Szklarczyk D, Kirsch R, Koutrouli M, Nastou K, Mehryary F, Hachilif R, Gable AL, Fang T, Doncheva NT, Pyysalo S, Bork P, Jensen LJ, von Mering C The STRING database in 2023: protein-protein association networks and functional enrichment analyses for any sequenced genome of interest. Nucleic Acids Res. 2023; 51(D1):D638-D646. doi: 10.1093/nar/gkac1000.PMID: 36370105
8. Huang da W, Sherman BT, Lempicki RA. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources Nat Protoc. 2009; 4(1): 44-57. doi: 10.1038/nprot.2008.211.PMID: 19131956
9. Ge SX, Jung D, Yao R. ShinyGO: a graphical gene-set enrichment tool for animals and plants Bioinformatics. 2020; 36(8): 2628-2629. doi: 10.1093/bioinformatics/btz931.PMID: 31882993 Free PMC article.
10. Pomaznoy, M., Ha, B. and Peters, B. GOnet: a tool for interactive Gene Ontology analysis. BMC Bioinformatics, 2018; 19: 470. https://doi.org/10.1186/s12859-018-2533-3
11. Mark Davies, Michał Nowotka, George Papadatos, Nathan Dedman, Anna Gaulton, Francis Atkinson, Louisa Bellis, John P. Overington ChEMBL web services: streamlining access to drug discovery data and utilities Nucleic Acids Research, Volume 43, Issue W1, 2015, Pages W612-W620, https://doi.org/10.1093/nar/gkv352
12. Keith J Kelleher, Timothy K Sheils, Stephen L Mathias, Jeremy J Yang, Vincent T Metzger, Vishal B Siramshetty, Dac-Trung Nguyen, Lars Juhl Jensen, Dušica Vidović, Stephan C Schürer Pharos 2023: an integrated resource for the understudied human proteome Nucleic Acids Research, Volume 51, Issue D1, 2023; Pages D1405-D1416: https://doi.org/10.1093/nar/gkac1033
13. Zhang M, Ma J, Guo Q, Ding S, Wang Y, Pu H. CD8+ T Cell-Associated Gene Signature Correlates with Prognosis Risk and Immunotherapy Response in Patients with Lung Adenocarcinoma. Front Immunol. 2022; 13: 806877. doi: 10.3389/fimmu.2022.806877. PMID: 35273597; PMCID: PMC8902308.
14. Xiao K, Xiao K, Li K, Xue P, Zhu S. Prognostic Role of TIGIT Expression in Patients with Solid Tumors: A Meta-Analysis. J Immunol Res. 2021; 2021: 5440572. doi: 10.1155/2021/5440572. PMID: 34888386; PMCID: PMC8651431
15. Punt S, Houwing-Duistermaat JJ, Schulkens IA, Thijssen VL, *et al.*, Correlations between immune response and vascularization qRT-PCR gene expression clusters in squamous cervical cancer. Mol Cancer. 2015; 14: 71. doi: 10.1186/s12943-015-0350-0. PMID: 25889974; PMCID: PMC4392729.
16. Li YW, Dou BM, Wang SJ, Lü ZX *et al.*, Acupuncture alleviates inflammatory pain by activating CXCL1/CXCR2 signaling in the primary somatosensory cortex in adjuvant induced arthritis rats. Zhen Ci Yan Jiu. 2024; 49(4): 331-40. English, Chinese. doi: 10.13702/j.1000-0607.20230107. PMID: 38649200.
17. Li LL, Zhou DX, Lu M, Zhou D, *et al.*, An integrated biomarker of PD-L1 expression and intraepithelial CD8+ T cell infiltration was associated with the prognosis of lung cancer patients after intracranial resection of brain metastases. Thorac Cancer. 2022; 13(13): 1948-1960. doi: 10.1111/1759-7714.14473. Epub 2022 May 20. PMID: 35596539; PMCID: PMC925083
18. Yang Z, Peng Y, Xu J, Chen P, *et al.*, PVR/TIGIT and PD-L1/PD-1 expression predicts survival and enlightens combined immunotherapy in lung squamous cell carcinoma. Transl Oncol. 2022; 24: 101501. doi: 10.1016/j.tranon.2022.101501. Epub 2022 Aug 1. PMID: 35926369; PMCID: PMC9352965
19. Tagliabue M, Maffini F, Fumagalli C, Gandini S, *et al.*, A role for the immune system in advanced laryngeal cancer. Sci Rep. 2020; 10(1): 18327. doi: 10.1038/s41598-020-73747-0. Erratum in: Sci Rep. 2021; 11(1): 9760. doi: 10.1038/s41598-021-89174-8. PMID: 33110100; PMCID: PMC7591515.
20. Wang P, Han L, Yu M, Cao Z, Li X, Shao Y, Zhu G. The Prognostic Value of PERK in Cancer and Its Relationship With Immune Cell Infiltration. Front Mol Biosci. 2021; 8: 648752. doi: 10.3389/fmolb.2021.648752. PMID: 33937330; PMCID: PMC8085429.
21. West NR, Milne K, Truong PT, *et al.*, Tumor-infiltrating lymphocytes predict response to anthracycline-based chemotherapy in estrogen receptor-negative breast cancer. Breast Cancer Res. 2011; 13(6): R126. doi: 10.1186/bcr3072. Epub 2011 Dec 8. PMID: 22151962; PMCID: PMC3326568.
22. Xiao Y, Li ZZ, Zhong NN, Cao LM, Liu B, Bu LL. Charting new frontiers: Co-inhibitory immune checkpoint proteins in therapeutics, biomarkers, and drug delivery systems in cancer care. Transl Oncol. 2023; 38: 101794. doi: 10.1016/j.tranon.2023.101794. Epub 2023 Oct 9. PMID: 37820473; PMCID: PMC10582482.
23. Wu D, Wang X. Application of clinical bioinformatics in lung cancer-specific biomarkers. Cancer Metastasis Rev. 2015; 34(2): 209-16. doi: 10.1007/s10555-015-9564-2. PMID: 25933723.

Cite this article: Anbuselvi S, Saranya S. Identification of Hub Genes and Therapeutic Targets in Lung Cancer Using Software Tools. Asian J Biol Life Sci. 2025;14(2):400-9.