EUROPEAN JOURNAL OF BIOMEDICAL AND PHARMACEUTICAL SCIENCES
http://www.ejbps.com

ISSN 2349-8870
Coden USA: EJBPBO
Volume: 13
Issue: 7
Year: 2026

IDENTIFICATION OF THERAPEUTIC TARGETS AND PROGNOSTIC BIOMARKERS THROUGH INTEGRATIVE DRUG–GENE INTERACTION AND SURVIVAL ANALYSIS OF LUNG CANCER HUB GENES

Firdos Khanam[1], Muchokota Sushma*[2]

[1]Pharm D PB 2nd Year, Aditya Bangalore Institute of Pharmacy Education and Research (ABIPER), Kogilu Main Road, Yelahanka, Bangalore-560064, Karnataka, India.
[2*]Pharm D, Ph. D (Pharmacy Practice), Associate Professor, Department of Pharmacy Practice, Aditya Bangalore Institute of Pharmacy Education and Research (ABIPER), Kogilu Main Road, Yelahanka, Bangalore-560064, Karnataka, India.

***Corresponding Author: Muchokota Sushma**

Pharm D, Ph. D (Pharmacy Practice), Associate Professor, Department of Pharmacy Practice, Aditya Bangalore Institute of Pharmacy Education and Research (ABIPER), Kogilu Main Road, Yelahanka, Bangalore-560064, Karnataka, India.

DOI: https://doi.org/10.5281/zenodo.21290215

How to cite this Article: Firdos Khanam[1], Muchokota Sushma*[2] (2026). Identification Of Therapeutic Targets And Prognostic Biomarkers Through Integrative Drug–Gene Interaction And Survival Analysis Of Lung Cancer Hub Genes. European Journal of Biomedical and Pharmaceutical Sciences, 13(7), 315–325.

Article Received on 13/06/2026 **Article Revised on 03/07/2026** **Article Published on 10/07/2026**

ABSTRACT

Background: Lung cancer is still a major public health problem in the world, and it is still one of the leading causes of cancer mortality. Non-small cell lung cancer (NSCLC), which mainly includes lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), constitutes the majority of diagnosed lung cancer cases. While significant advances have been made in treatment strategies, it is still crucial to identify potentially useful molecular markers and clinically relevant therapeutic targets to advance patient care and disease management. **Methods:** An integrated bioinformatics platform was used to analyze transcriptomic, clinical, and pharmacogenomic data from publicly available databases, such as The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), STRING, DrugBank, and DGIdb. Differential expression analysis was used to find genes related to NSCLC. Protein–protein interaction network construction, functional enrichment analysis, prognostic evaluation, and drug–gene interaction analysis were then conducted to identify the biological and clinical significance of the identified genes. **Results:** We identified several candidate genes that could be of diagnostic and therapeutic value in NSCLC. EGFR, KRAS, MET, TP53, and PIK3CA were identified as the major hub genes of the most crucial cancer-related pathways in LUAD. The genes identified as central regulatory genes associated with cell-cycle control and tumour progression in LUSC were TOP2A, CCNB1, CCNA2, and BUB1B. Survival analyses showed that strong relationships exist between expression levels of these genes and overall patient survival. Additionally, drug–gene interaction analysis revealed several approved and Phase II/III drugs that could exploit these molecular abnormalities. **Conclusion:** This study proves that the integration of transcriptomic and pharmacogenomic data, in conjunction with the system biology approach, offers valuable insights into the molecular features of NSCLC.

KEYWORDS: Lung Cancer, Non-Small Cell Lung Cancer (NSCLC), Lung Adenocarcinoma (LUAD), Lung Squamous Cell Carcinoma (LUSC), Protein–Protein Interaction Network and Drug–Gene Interaction.

INTRODUCTION

Lung cancer is one of the most prevalent cancers in the world, and is also the leading cause of cancer-related deaths, accounting for a significant proportion of cancer deaths every year.[1] Yet, the prognosis for many remains poor due to the lack of early diagnosis and treatment, disease heterogeneity, resistance, and frequent relapses.[2] Nearly 85% of all lung cancer cases are non-small cell lung cancer (NSCLC), which is further subclassified into lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), and has different molecular and clinicopathological characteristics that affect the

progression and response to treatment.[3] The recent advances in high-throughput sequencing and computational biology have allowed in-depth characterization of the genomic, transcriptomic, and proteomic changes that play a role in the development of lung cancer.[4]

A combination of large-scale publicly accessible datasets is an effective tool for the identification of differentially expressed genes (DEGs), altered signalling pathways, and molecular networks related to tumour progression. These analyses are useful in the discovery of biomarkers that could potentially be used in diagnosis, prognosis, and precision-based treatment strategies.[5] Protein–protein interaction (PPI) network analysis is one of the most widely used methods to predict hub genes, which are highly central genes and play roles in the control of key cellular processes. These genes tend to be related to cell proliferation, apoptosis, angiogenesis, invasion, and metastasis, which all play a role in cancer progression.[6] Earlier studies have shown that genes like EGFR, KRAS, TP53, MET, and PIK3CA are significantly involved in lung tumorigenesis and affect clinical outcomes and targeted therapy sensitivity.[7-8]

Thus, systematic identification of hub genes could aid in understanding disease mechanisms and allow for the identification of clinically relevant therapeutic targets. Finally, survival analysis is an important tool for assessing the clinical significance of candidate genes, as it is possible to assess associations between gene expression and patient outcome.[9] Genes that are associated with positive or negative survival trends can be used as biomarkers of prognosis and help stratify patients and guide treatment choices. Integrating survival information with molecular findings further enhances the translational significance of identified targets.[10] Drug–gene interaction analysis has become an important component of precision oncology because it enables the identification of therapeutic agents capable of targeting key molecular alterations in cancer.[11]

By integrating hub genes with pharmacological databases, clinically actionable targets can be identified, and opportunities for personalized treatment approaches can be explored. In addition, this strategy supports drug repurposing efforts, potentially reducing the cost and time required for the development of new anticancer therapies.[12] In the context of this study, an integrated bioinformatics approach was used to select molecular targets for lung cancer therapy, which involved differential gene expression analysis, functional enrichment, construction of protein–protein interaction networks, survival analysis, and drug–gene interaction analysis.

OBJECTIVES

- To identify genes showing significant expression differences between lung cancer and normal tissues through the integrated evaluation of transcriptomic data obtained from the TCGA and GEO databases.
- To characterize interactions among differentially expressed genes by constructing protein–protein interaction networks and performing functional enrichment analyses in order to determine key regulatory genes and biological pathways associated with lung cancer progression.
- To assess the therapeutic potential of selected hub genes by investigating their relationships with approved and emerging pharmacological agents and identifying opportunities for targeted therapy and drug repurposing.
- To evaluate the association between candidate gene expression profiles and clinical outcomes, particularly overall survival, and determine their usefulness as prognostic markers in personalized cancer care.
- To confirm the expression patterns and biological relevance of prioritized genes using independent datasets and major non-small cell lung cancer subtypes, including lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC), with the aim of identifying shared as well as subtype-specific therapeutic targets.

METHODOLOGY

Study Selection

Combination of various bioinformatics resources was used to investigate the molecular changes and potential therapeutic targets in lung cancer in the present study. Prior to analysis, the literature regarding the genes associated with lung cancer, expression signatures, molecular pathways and therapeutic agents was compiled, using scientific databases such as Google Scholar and PubMed. The results from the previous studies were helpful in selecting the right data, right analytical tools and methodology for the study.

Source of Data

The data for this research were from publicly available databases. Data on gene expression, clinical data and survival information were obtained from The Cancer Genome Atlas (TCGA). To validate the data sets were independently obtained from Gene Expression Omnibus (GEO). cBioPortal and the UCSC Xena Browser were used to gain further genomic information. Protein interactions data were downloaded from STRING and Drug–Target relationships data were downloaded from DrugBank and Drug–Gene Interaction Database (DGIdb). These resources comprised transcriptomic, clinical and pharmacological data that were required for analysis.

Study Design

This study was performed as a retrospective computational study using the existing public data. No recruitment of patients or any clinical intervention was performed, the information had been obtained and was available in online repositories. Special attention was

given to the identification of the differentially expressed genes, their biological relevance, and their therapeutic potential in lung cancer.

Study Period
The period of the work was from November 2025 to April 2026.

Study Site
The study used only publicly available databases and analytical platforms on the web, and did not involve any particular institution or physical location.

Study Population
Analysis was conducted on patients who were diagnosed with non-small cell lung cancer (NSCLC), with a focus on lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC). Samples from these subtypes with study criteria were chosen for further evaluation.

Sample Size
All available, eligible samples were included across all the selected TCGA cohorts. The data set had about 500–600 LUAD patients and 450–550 LUSC patients, which made the total count of patients around 1,000–1,100. The sample size was not calculated, as a secondary data analysis was conducted.

Eligibility Criteria
Inclusion Criteria
- Samples belonging to the TCGA-LUAD or TCGA-LUSC cohorts.
- Availability of complete gene expression data.
- Primary tumor tissue samples.
- Availability of corresponding clinical and survival information.

Exclusion Criteria
- Incomplete clinical records.
- Absence of survival information.
- Duplicate samples.
- Poor-quality or missing gene expression data.

Data Collection
RNA sequencing datasets were downloaded from the chosen databases to obtain information for gene expression. Clinical parameters like age, sex, tumour stage, and the outcomes of survival were gathered wherever possible. Downstream analyses also were performed to obtain information related to the biological pathways, protein interactions, and drug–gene associations.

Data Processing
The downloaded datasets have been pre-processed for the enhancement of data quality and consistency. Technical variations were minimized using normalisation procedures. Samples with very low expression for a given gene were not included in further analyses. Quality assessment was also carried out to identify and eliminate data that were unreliable data.

Differential Expression Analysis
Differential expression analysis identified genes with differences in expression between cancerous and normal tissues. The significant genes are selected using absolute log2 fold change ≥ 1 and false discovery rate (FDR) <0.05. The following analyses were performed on these genes.

Survival Analysis
Survival analysis was used to evaluate the correlation between gene expression and patient survival. Survival probabilities were plotted using Kaplan–Meier curves between groups of varying expression levels. Using Cox proportional hazards models, the impact of gene expression on overall survival was estimated.

Protein–Protein Interaction Analysis
Protein–protein interaction networks for the identified genes were built using information from the STRING database to investigate functional relationships. The Cytoscape software was used to visualize and analyse the network. The hub genes were identified using the topological characteristics of the genes in the network.

Functional Enrichment Analysis
To understand the biological functions of the genes identified, the Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses were performed. The results from these analyses included information related to biological processes, molecular functions, cellular components, and signalling pathways involved in lung cancer.

Drug–Gene Interaction Analysis
Selected hub genes were analysed for their therapeutic relevance by DrugBank and DGIdb. Known pharmacological agents were evaluated for their interactions with genes to find approved drugs, investigational compounds, and potential candidates for drug repurposing.

External Validation
Independent GEO datasets were used to confirm the results of the primary analysis. To check the consistency of the results, expression patterns and clinical significance of the selected genes were analysed in various patient groups and in the various subtypes of NSCLC.

Study Workflow
Several analytical steps were performed in the study, including data input and preprocessing, differential expression analysis, network construction, functional enrichment analysis and lastly survival assessment, drug–gene interaction evaluation and final validation with external data sets. Using this workflow, it was possible to identify those genes which are important in

the biology of lung cancer, and other potential therapeutic targets identified.

Statistical Analysis

All analyses were done in R software. For gene expression, survival, and enrichment analyses, packages were used that included DESeq2, limma, survival, and clusterProfiler. Differential expression of genes was performed considering the absolute log2 fold change ≥1 and FDR <0.05. Kaplan–Meier and Cox regression methods were used to evaluate survival outcomes. P value < 0.05 was considered statistically significant.

RESULTS

The appropriate datasets were retrieved from the Gene Expression Omnibus (GEO) repository, assessed for quality, and pre-processed before analysis. A differential expression analysis was then undertaken to determine the genes that have significant differences between the expression of cancer versus normal tissue samples. These expression changes were plotted in volcano plots and heatmap visualizations to provide a distribution and size representation. Further analyses were performed to gain insight into the biological importance of the discovered genes. Biological processes and signalling pathways most strongly correlated with the observed changes in gene expression were determined using functional enrichment methods. Interaction networks were also created to analyse the relationships between all the proteins that were encoded by the differentially expressed genes, with the aim of finding genes that play key roles. Beyond this, drug–gene interaction analysis was performed to assess the therapeutic applications of selected targets and to discover molecules with potential clinical applications. The results of LUAD and LUSC analyses are presented in the next sections.

Lung adenocarcinoma

To explore possible therapeutic applications in lung adenocarcinoma (LUAD), the candidate genes identified were assessed for their potential using information in the Drug–Gene Interaction Database (DGIdb). The goal of this analysis was to identify such genes that might be available for pharmacological targeting or as targets for new drug development. A number of the genes evaluated showed some documented interactions with approved drugs and drugs in development. These interactions are particularly interesting because they may have clinical significance and could be the basis for developing specific treatment strategies. In conclusion, the selected genes could be useful molecular targets for targeted treatment in patients with LUAD and offer additional support to the current findings.

Table 1: Drug–Gene Interactions Identified in LUAD.

Gene	Drug Name	Approval Status	Interaction Type	Source
MET	Cabozantinib	Not Approved	Inhibitor	DGIdb
MET	Damufetinib	Not Approved	Inhibitor	DGIdb
MET	Tyrphostin AG	Not Approved	Inhibitor	DGIdb
MET	AMG-337	Not Approved	Inhibitor	DGIdb
MET	Crizotinib	Approved	Inhibitor	DGIdb
MET	Capmatinib	Approved	Inhibitor	DGIdb
MET	Golvatinib	Not Approved	Inhibitor	DGIdb
MET	Savolitinib	Not Approved	Inhibitor	DGIdb
MET	SGX523	Not Approved	Kinase Inhibitor	DGIdb

Table 1 & Fig. 1 results indicate that several genes, particularly MET, are associated with multiple therapeutic agents, highlighting their potential as druggable targets in LUAD.

Term	Gene	Drug	Regulatory Approval	Indication	Interaction Score
MET	MET	MK-2461	Not Approved		0.06
MET	MET	PF-04217903	Not Approved		0.66
MET	MET	CABOZANTINIB S-MALATE	Approved	antineoplastic agent	0.27
MET	MET	CAPMATINIB	Approved		1.5
MET	MET	GLESATINIB	Not Approved	antineoplastic agent	0.25
MET	MET	CRIZOTINIB	Approved	antineoplastic agent	0.4
MET	MET	TEPOTINIB	Approved		0.87

Figure 1: DGIdb drug–gene interaction network for LUAD.

Table 2: Drug–Gene Interaction Analysis of LUAD.

Gene	No. of Drugs
TP53	452
EGFR	192
PIK3CA	215
MET	104

Among the identified genes, TP53 had the greatest number of drug interactions, followed by PIK3CA, EGFR, and MET, suggesting their high therapeutic potential in LUAD (Table 2).

Prognostic Validation and Precision Oncology Application

Prognostic Analysis and Survival Evaluation of LUAD Hub Genes

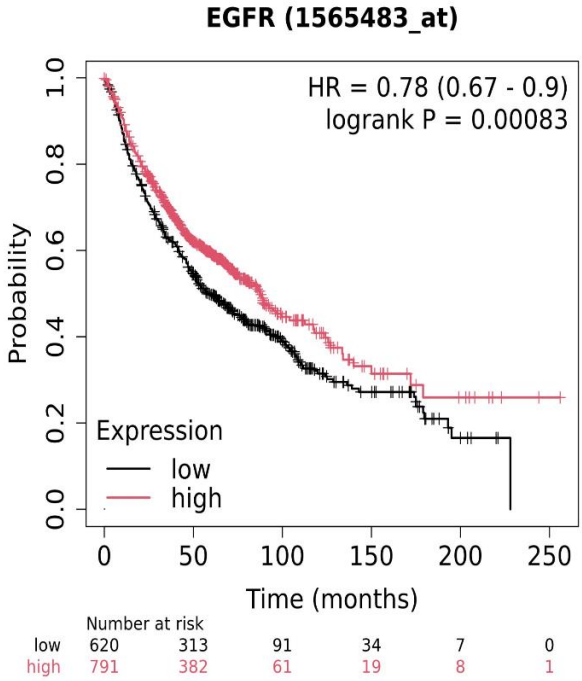

Figure 2: Kaplan–Meier survival curve of EGFR.

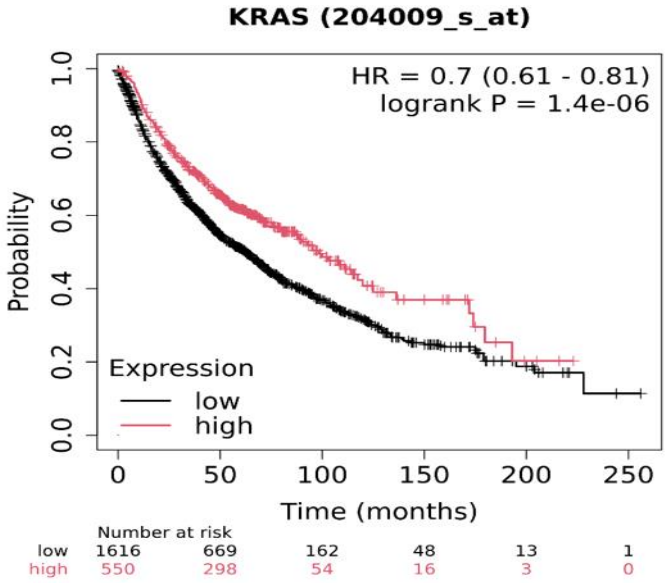

Figure 3: Kaplan–Meier survival curve of KRAS.

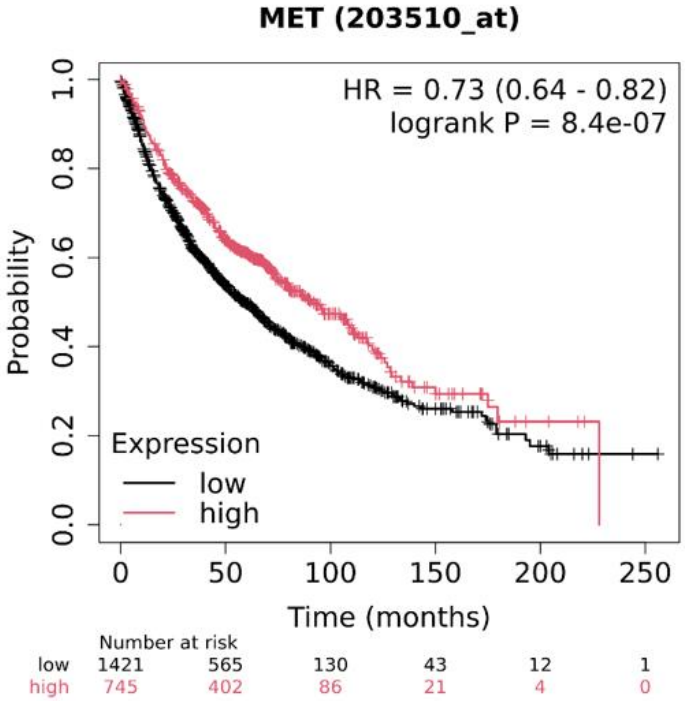

Figure 4: Kaplan–Meier survival curve of MET.

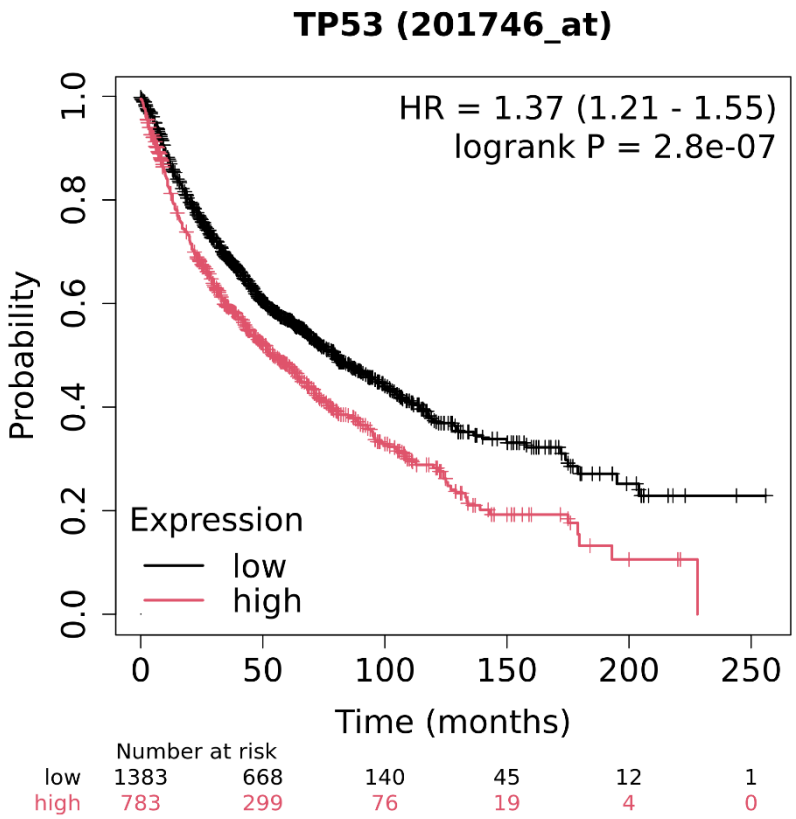

Figure 5: Kaplan–Meier survival curve of TP53.

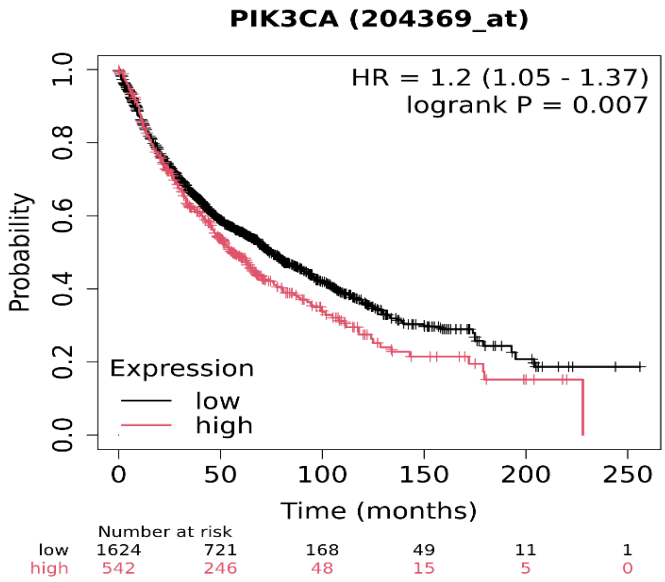

Figure 6: Kaplan–Meier survival curve of PIK3CA.

Table 3: Prognostic significance of selected LUAD hub genes based on Kaplan–Meier survival analysis.

Gene	Affy ID	p-value	FDR	Median Survival (Low Expression)	Median Survival (High Expression)	Prognostic Significance
EGFR	1565483_at	0.00083	Significant	59	88	Favorable
KRAS	204009_s_at	1.4×10^{-6}	1%	62.3	98.5	Favorable
MET	203510_at	8.4×10^{-7}	1%	59	91	Favorable
TP53	201746_at	2.8×10^{-7}	1%	79.87	54.3	Poor
PIK3CA	204369_at	0.007	>50%	74	56	Poor

The results of the Kaplan–Meier survival analysis used to assess the prognostic value of selected hub genes in lung adenocarcinoma are shown in Figures 2-6 and Table 3. There was a significant correlation between the expression pattern of the genes and overall survival, as demonstrated in the analysis. Higher expression of EGFR, KRAS, and MET was associated with a favourable prognostic role with a better median survival. The expression of TP53 and PIK3CA, in contrast, was linked to poorer survival, suggesting a possible impact on clinical outcome. Genes that were analysed with the highest level of statistical significance are useful to predict the outcome of the disease, such as the genes of KRAS, MET, and TP53. Based on these observations, the hub genes identified in this study could be used as useful prognostic markers; they might be used for risk stratification, predicting prognosis for LUAD patients, and creating personalized therapeutic strategies for patients with LUAD.

Clinical and Precision Oncology Implications of LUAD Hub Genes

Several molecular targets of clinical relevance were identified in lung adenocarcinoma for the integrated evaluation of differential gene expression and protein–protein interaction networks, drug–gene interactions, and survival data. The main hub genes were EGFR, MET, TP53, PIK3CA and KRAS, and they were deemed to be of significant biological relevance and translation potential. These genes are thus active in key signalling pathways which control cell growth, survival, differentiation and cancer progression. The drug–gene interaction analysis revealed that EGFR and MET are related to a number of approved targeted drugs that are already used in the clinic, confirming that they have therapeutic relevance. Moreover, PIK3CA was identified as a promising target for future precision medicine therapies. Survival analyses showed that these genes were significantly correlated with clinical outcomes, suggesting their potential as prognostic markers. All these results point to the possible use of these hub genes in patient stratification, treatment selection, and ongoing development of precision oncology approaches for LUAD.

Lung squamous cell carcinoma Analysis

Survival Analysis and Prognostic Evaluation of LUSC Hub Genes

CDK1

The selected database did not contain Kaplan–Meier survival data for the corresponding CDK1 gene; therefore, survival analysis was not possible, and the selected gene was initially thought to be a potential hub gene. For those genes for which survival data were available, prognostic evaluation was therefore carried out for the genes TOP2A, CCNB1, CCNA2, and BUB1B.

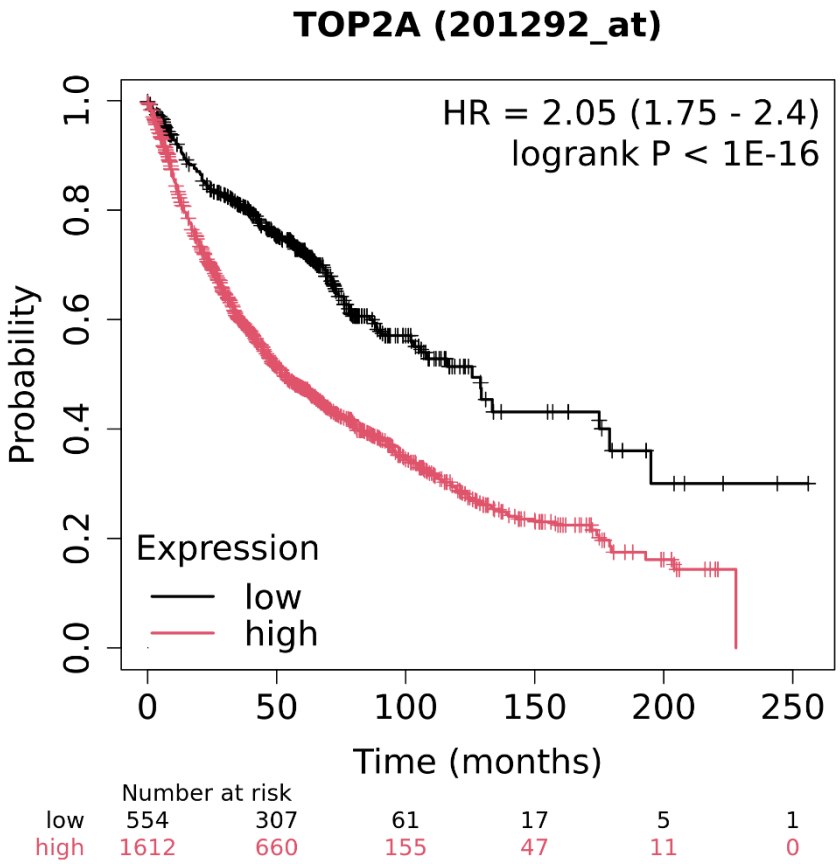

Figure 7: Kaplan–Meier survival curve of TOP2A.

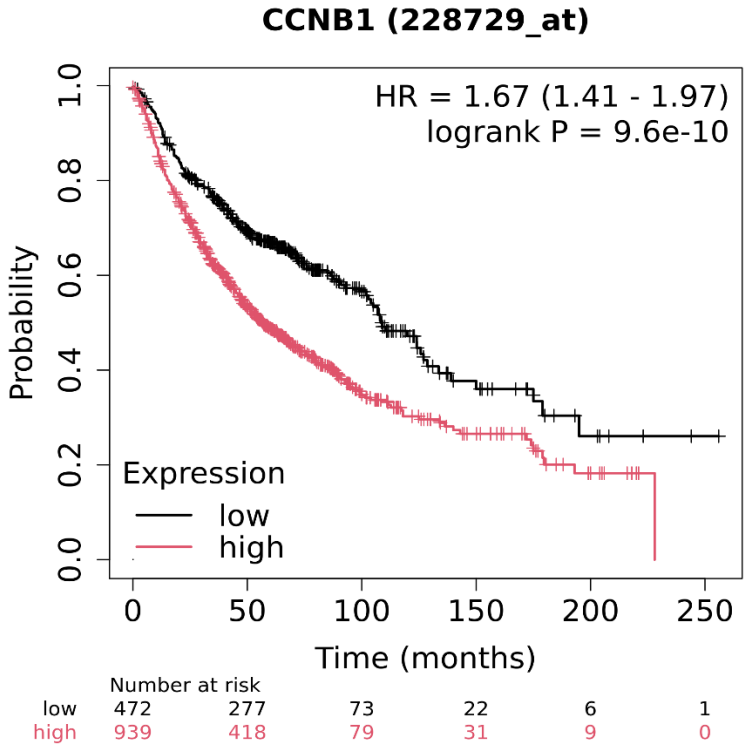

Figure 8: Kaplan–Meier survival curve of CCNB1.

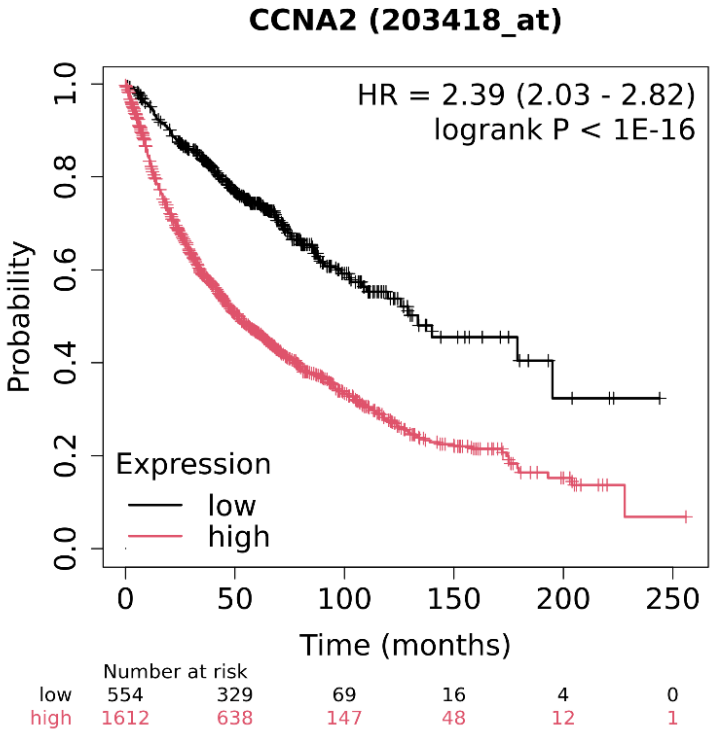

Figure 9: Kaplan–Meier survival curve of CCNA2.

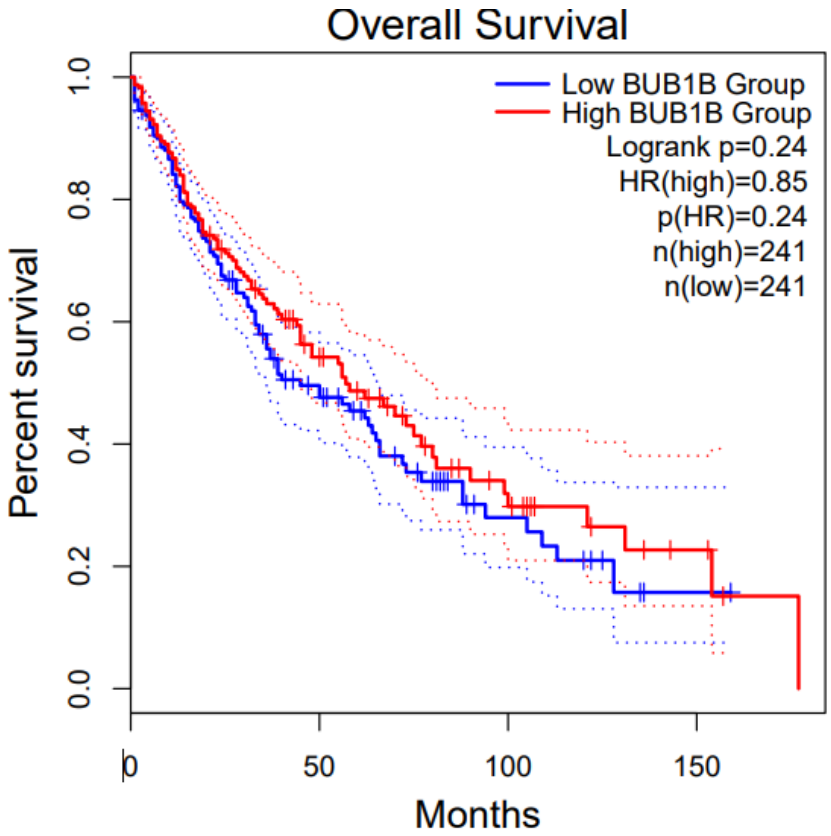

Figure 10: Kaplan–Meier survival curve of BUB1B.

CDK1 was not analysed by survival analysis because the probe wasn't available in the chosen survival analysis platform. Therefore, other prioritized hub genes identified in LUSC were used to perform prognostic assessment.

Table 4: Prognostic significance of selected LUSC hub genes based on Kaplan–Meier survival analysis.

Gene	Affy ID	p-value	FDR	Median Survival (Low Expression)	Median Survival (High Expression)	Prognostic Significance
TOP2A	201292_at	$<1 \times 10^{-16}$	1%	125.77	52	Poor
CCNB1	228729_at	9.6×10^{-10}	1%	108.97	57	Poor
CCNA2	203418_at	$<1 \times 10^{-16}$	1%	133.57	52	Poor
BUB1B	203755_at	$<1 \times 10^{-16}$	1%	116	52	Poor

Figures 7–10 and Table 4 show the results of the Kaplan–Meier survival analysis used to assess the prognostic value of selected hub genes in lung squamous cell carcinoma. The results revealed significant associations between TOP2A, CCNB1, CCNA2, and BUB1B and overall survival. Patients with high levels of these genes had a significantly shorter lifespan than those with low levels of the genes, suggesting that these genes are linked to a poor prognosis. In all, among the genes analysed, TOP2A, CCNA2, and BUB1B had very high statistical significance ($p < 1 \times 10^{-16}$), and CCNB1 was also significantly associated with survival ($p = 9.6 \times 10^{-10}$). These findings indicate that regulation of cell cycle pathways, including mitotic checkpoint and cell cycle activity, could be significant in LUSC tumour progression and poor clinical prognosis. So, these genes can be potential indicators for prognosis and can be used for patient stratification and individualized treatment strategies.

Precision Oncology Application of LUSC Hub Genes

Integrated analysis of differential gene expression, protein–protein interaction networks, drug–gene interactions, and survival showed that the molecules TOP2A, CCNB1, CCNA2, and BUB1B are important molecular drivers in LSCC. These genes are critical for other cell functions, including cell cycle control, the control of the mitotic checkpoint, and the segregation of chromosomes, many of which are altered in the initiation and progression of tumours. The expression of these is raised, and there is a strong correlation with overall survival, suggesting their possible use as markers of prognosis and disease severity. In addition, drug–gene interaction analysis revealed a number of therapeutic possibilities that were associated with pathways related to cell-cycle regulation. In summary, these findings highlight the clinical relevance of the LINC complex members (TOP2A, CCNB1, and CCNA2) and BUB1B in the prognosis of LUSC and targeted therapeutic interventions based on precision medicine, thus helping the development of precision medicine–based treatment strategies.

DISCUSSION

Despite significant advances in the diagnosis and treatment of lung cancer, the disease is still one of the top ten leading causes of cancer-related mortality worldwide, partly because it is often diagnosed late in its progression. In addition to this, there is a considerable molecular diversity in lung cancer, and responses to the available treatment modalities are variable. In the last few years, the study of lung tumorigenesis has become more molecular and complex, owing to the recent developments in genomics, computational biology, and systems-level analyses that have allowed deeper insight into molecular events that drive lung tumorigenesis and the discovery of clinically important therapeutic targets. To investigate molecular changes and targetable therapeutic options for lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) in the present study, an integrated framework composed of transcriptomic analysis, network inference, and pharmacogenomics was used.

The gene expression patterns of the two main subtypes of non-small cell lung cancer (NSCLC) were found to be different by differential expression analysis. In the selected statistical thresholds, LUAD presented a comparatively smaller number of DE genes, mostly decreased in expression level, which were associated with the loss of genes related to normal cellular functions. LUSC, in contrast, had a broader distribution of both overexpression and underexpression of genes, suggesting a more significant level of transcript dysregulation. These results align with previous findings that LUAD and LUSC have distinct molecular features at the level of two distinct diseases. Protein–protein interaction analysis revealed some highly connected genes potentially playing significant roles in the disease process.

EGFR, KRAS, TP53, MET, and PIK3CA were found to be focal nodes in the interaction network in LUAD. These genes have been demonstrated to be involved in signalling pathways that regulate invasion and metastatic behaviour, cellular survival, and proliferation. Mutations in the genes that code for EGFR and KRAS are known to be oncogenic events, and mutations in the TP53 are linked to defects in genomic surveillance. Similarly, MET and PIK3CA have been associated with increased aggressive tumour behaviour and decreased responsiveness to therapy. They also hold high positions in their network, which further substantiates their biological and clinical significance in LUAD. A different molecular pattern was observed in LUSC. The major hub genes identified included CDK1, TOP2A, CCNB1, CCNA2, BRCA1, BUB1B and AURKA. These genes are primarily involved in cell-cycle regulation, chromosome dynamics, mitotic control, and DNA repair mechanisms. Gene Ontology enrichment analysis reinforced the importance of biological processes related to cell division and chromosome maintenance.

Disruption of these functions can result in genomic instability and uncontrolled cellular growth, both of which are hallmarks of cancer development. Collectively, these observations suggest that abnormalities in cell-cycle regulatory pathways may have a central role in the pathogenesis of LUSC. KEGG pathway enrichment analysis did not identify statistically significant pathways in either subtype when the selected filtering criteria were applied. This finding may be attributable to strict selection thresholds, differences among datasets, or the limited number of significant genes identified in LUAD. Nevertheless, the biological functions and network characteristics of the prioritized genes indicate the involvement of several cancer-associated molecular processes. Drug–gene interaction analysis identified a number of therapeutically relevant targets. Several approved and investigational agents demonstrated interactions with the key genes identified in this study. Interestingly, drugs targeting EGFR (Gefitinib, Erlotinib, and Afatinib) had significant clinical relevance. Likewise, MET pathway inhibitors, e.g., Crizotinib and Capmatinib, were found to be therapeutic candidates of interest. Other specific drugs under investigation were the KRAS inhibitor Sotorasib and PI3K inhibitor Alpelisib.

Among the genes assessed, TP53 had the most reported drug interactions, indicating the high research interest in targeting molecular pathways related to TP53. Different biological types of the two NSCLCs were evident. Receptor tyrosine kinase signaling pathways through EGFR, KRAS, and MET seemed to be primarily responsible for the driving of LUAD. On the other hand, genes involved in cell-cycle progression and mitosis were more associated with LUSC. However, shared molecular targets like TP53, EGFR, PIK3CA, and MET were found in both subtypes, suggesting overlapping pathogenic mechanisms.

CONCLUSION

The present study demonstrated that the method of bioinformatics, systems biology and pharmacogenomic analysis can be used as a valuable tool in identifying the key molecular signatures, prognosis markers and therapeutic targets in non-small-cell lung cancer (NSCLC) including LUAD and LUSC. Molecular properties and clinically relevant targets for each of the subtypes (e.g., EGFR, KRAS, MET, TP53, and PIK3CA) were identified through the use of differential gene expression and functional enrichment analysis, protein–protein interaction networks, survival evaluation, and associations with drugs. A few agents were discovered that were approved, suggesting opportunities for precision and repurposing approaches. These findings combined can contribute to a deeper understanding of the molecular complexity of NSCLC and as a basis for more individual treatments and future translational research in lung cancer.

ACKNOWLEDGEMENT: We would like to thank your professors and principal.

CONFLICT OF INTEREST: Nil.

REFERENCES

1. Global Cancer Statistics 2024 Bray F, Laversanne M, Sung H, et al. Global cancer statistics 2024: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.
2. Herbst RS, Morgensztern D, Boshoff C. The biology and management of non-small cell lung cancer. Nature, 2018; 553: 446–454.
3. Duma N, Santana-Davila R, Molina JR. Non-small cell lung cancer: epidemiology, screening, diagnosis and treatment. Mayo Clin Proc, 2019; 94(8): 1623–1640.
4. Hanahan D. Hallmarks of cancer: new dimensions. Cancer Discov, 2022; 12(1): 31–46.
5. Clough E, Barrett T. The Gene Expression Omnibus database. Methods Mol Biol, 2016; 1418: 93–110.
6. Szklarczyk D, Gable AL, Lyon D, et al. STRING v11: protein–protein association networks. Nucleic Acids Res, 2019; 47: D607–D613.
7. Pao W, Girard N. New driver mutations in non-small-cell lung cancer. Lancet Oncol, 2011; 12(2): 175–180.
8. Skoulidis F, Heymach JV. Co-occurring genomic alterations in non-small-cell lung cancer biology and therapy. Nat Rev Cancer, 2019; 19(9): 495–509.
9. Győrffy B, Surowiak P, Budczies J, Lánczky A. Online survival analysis tools for cancer research. Breast Cancer Res Treat, 2013; 123: 725–731.
10. Tang Z, Li C, Kang B, Gao G, Li C, Zhang Z. GEPIA: a web server for cancer and normal gene expression profiling and survival analysis. Nucleic Acids Res, 2017; 45: W98–W102.
11. Freshour SL, Kiwala S, Cotto KC, et al. Integration of the Drug–Gene Interaction Database (DGIdb 4.0). Nucleic Acids Res, 2021; 49: D1144–D1151.
12. Pushpakom S, Iorio F, Eyers PA, et al. Drug repurposing: progress, challenges and recommendations. Nat Rev Drug Discov, 2019; 18(1): 41–58.