



OPEN ACCESS

EDITED BY

Eyad Elkord,
University of Salford, United Kingdom

REVIEWED BY

Galal Metwally,
Zagazig University, Egypt
Minglei Zhuo,
Peking University, China

*CORRESPONDENCE

Pengxia Liu

✉ liupengxia@imu.edu.cn

Changshan Wang

✉ changshanwang@imu.edu.cn

†These authors have contributed equally to this work

RECEIVED 07 June 2023

ACCEPTED 12 September 2023

PUBLISHED 29 September 2023

CITATION

Wu R, Ma R, Duan X, Zhang J, Li K, Yu L, Zhang M, Liu P and Wang C (2023) Identification of specific prognostic markers for lung squamous cell carcinoma based on tumor progression, immune infiltration, and stem index. *Front. Immunol.* 14:1236444. doi: 10.3389/fimmu.2023.1236444

COPYRIGHT

© 2023 Wu, Ma, Duan, Zhang, Li, Yu, Zhang, Liu and Wang. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Identification of specific prognostic markers for lung squamous cell carcinoma based on tumor progression, immune infiltration, and stem index

Rihan Wu^{1,2†}, Ru Ma^{1†}, Xiaojun Duan^{1,3†}, Jiandong Zhang¹, Kexin Li¹, Lei Yu¹, Mingyang Zhang¹, Pengxia Liu^{1*} and Changshan Wang^{1*}

¹School of Life Science, Inner Mongolia University, Hohhot, China, ²The Department of Oncology, Affiliated Hospital of Inner Mongolia Medical University, Hohhot, China, ³School of Basic Medicine, Inner Mongolia Medical University, Hohhot, China

Introduction: Lung squamous cell carcinoma (LUSC) is a unique subform of nonsmall cell lung cancer (NSCLC). The lack of specific driver genes as therapeutic targets leads to worse prognoses in patients with LUSC, even with chemotherapy, radiotherapy, or immune checkpoint inhibitors. Furthermore, research on the LUSC-specific prognosis genes is lacking. This study aimed to develop a comprehensive LUSC-specific differentially expressed genes (DEGs) signature for prognosis correlated with tumor progression, immune infiltration, and stem index.

Methods: RNA sequencing data for LUSC and lung adenocarcinoma (LUAD) were extracted from The Cancer Genome Atlas (TCGA) data portal, and DEGs analyses were conducted in TCGA-LUSC and TCGA-LUAD cohorts to identify specific DEGs associated with LUSC. Functional analysis and protein-protein interaction network were performed to annotate the roles of LUSC-specific DEGs and select the top 100 LUSC-specific DEGs. Univariate Cox regression and least absolute shrinkage and selection operator regression analyses were performed to select prognosis-related DEGs.

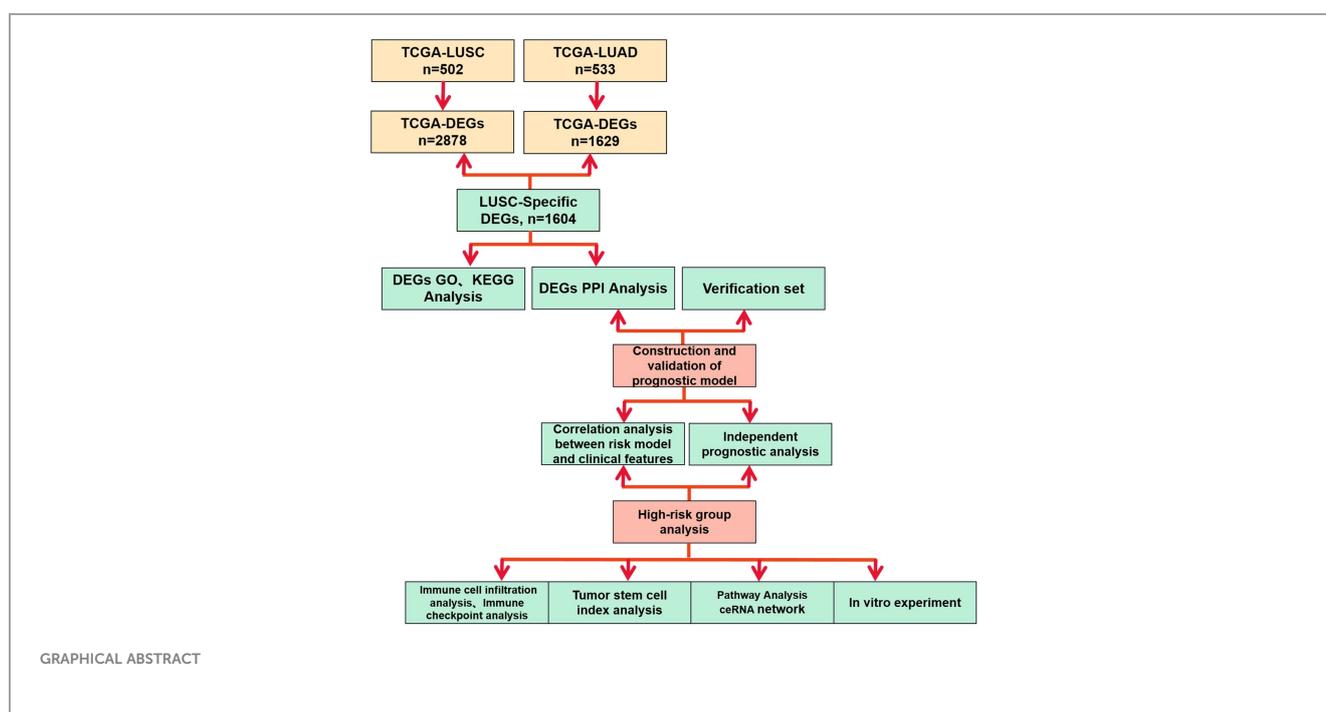
Results: Overall, 1,604 LUSC-specific DEGs were obtained, and a validated seven-gene signature was constructed comprising FGG, C3, FGA, JUN, CST3, CPSF4, and HIST1H2BH. FGG, C3, FGA, JUN, and CST3 were correlated with poor LUSC prognosis, whereas CPSF4 and HIST1H2BH were potential positive prognosis markers in patients with LUSC. Receiver operating characteristic analysis further confirmed that the genetic profile could accurately estimate the overall survival of LUSC patients. Analysis of immune infiltration demonstrated that the high risk (HR) LUSC patients exhibited accelerated tumor infiltration, relative to low risk (LR) LUSC patients. Molecular expressions of immune checkpoint genes differed significantly between the HR and LR cohorts. A ceRNA network containing 19 lncRNAs, 50 miRNAs, and 7 prognostic DEGs was constructed to demonstrate the prognostic value of novel biomarkers of LUSC-specific DEGs based on tumor progression, stemindex, and immune infiltration. In vitro experimental models confirmed

that LUSC-specific DEG FGG expression was significantly higher in tumor cells and correlated with immune tumor progression, immune infiltration, and stem index. *In vitro* experimental models confirmed that LUSC-specific DEG FGG expression was significantly higher in tumor cells and correlated with immune tumor progression, immune infiltration, and stem index.

Conclusion: Our study demonstrated the potential clinical implication of the 7-DEGs signature for prognosis prediction of LUSC patients based on tumor progression, immune infiltration, and stem index. And the FGG could be an independent prognostic biomarker of LUSC promoting cell proliferation, migration, invasion, THP-1 cell infiltration, and stem cell maintenance.

KEYWORDS

LUSC, prognosis, biomarker, tumor microenvironment, cancer stem cell



1 Introduction

Lung cancer is heterogeneous and fatal, with non-small cell lung cancer (NSCLC) as its main pathological subtype. Lung squamous cell carcinoma (LUSC) and lung adenocarcinoma (LUAD) are the primary subtypes of NSCLC (1); however, they differ in many aspects, including the origin of cells, genetic variation, epigenetics,

Abbreviations: EMT, Epithelial-Mesenchymal Transition; NSCLC, Non-small cell carcinoma; LUAD, Lung adenocarcinoma; LUSC, Lung squamous cell carcinoma; mRNA, Message ribonucleic acid; FGG, Fibrinogen gamma; GO, Gene Ontology; KEGG, Kyoto Encyclopedia of Genes and Genomes; PPI, Protein-protein interaction; TCGA, The Cancer Genome Atlas; OS, Overall survival.

and the outcome of antineoplastic drugs (2). Despite tremendous advances in diagnosis and treatment, including molecular targeted therapeutics and immunotherapy, the clinical outcomes of LUSC remain unsatisfactory (3). Patients with LUSC are often diagnosed in an advanced stage when existing therapy cannot be administered in a timely manner (4). LUSC patients are also not as sensitive as LUAD patients to chemotherapy, radiotherapy, and tumor immunotherapy. In addition, the prognosis of LUSC is poor, with an estimated 5-year survival rate of <15% (5). Therefore, distinguishing LUSC from LUAD is important to identify effective prognostic biomarkers.

Although studies based on the whole genome (6, 7), epigenetics (8), cancer stem cells (CSCs) (9, 10), and tumor microenvironment

(TME) (11) have analyzed differentially expressed genes (DEGs) in LUSC and LUAD, research on the LUSC-specific prognosis genes is lacking. A previous study (4), involving 178 LUSC cases, conducted using the Cancer Genome Atlas (TCGA) Research Network reported complex genomic alterations in LUSC, including significant copy number alterations, which peaked for SOX2, PDGFRA/KIT, EGFR, FGFR1, CCND1, and CDKN2A. In LUSC, CDKN2A/RB1, NFE2L2/KEAP1, squamous differentiation genes, and PI3K/Akt were significantly altered. TP53 is the most commonly mutated gene with a mutation frequency of > 80% in LUSC. The overexpression and amplification of genes, SOX2 and TP63, are spectrum factors of LUSC (12). Despite progress in research on biomarkers for LUSC, oncology targets are rare. Recent studies on genetic biomarkers for LUSC have focused on a single gene based on the cognitions of CSCs and the TME in cancer progression, as well as drug resistance and response to immune checkpoint blockade. Traditional methods use differential expression detection to identify potential biomarkers but may miss out on useful genes. As the occurrence and development of malignant tumors is a long-term complex process involving genomic changes, the interaction between tumor cells and their immune microenvironment, and the participation of tumor stem cells, the behavior cognition of malignant tumors warrants extensive research.

Therefore, in this study, we aimed to compare DEGs of LUSC with LUAD using biological information analytical methods based on prognostic risk factors, including tumor invasion, metastasis, survival, immune infiltration, and tumor stem cell-related genes. DEGs in LUSC were employed to generate a risk model to evaluate the prognostic value of characteristic genes for possible prognostic indicators or therapeutic targets for LUSC. We further explored the associations between the specific prognostic markers FGG and tumor progression, immune invasion, and the tumor cell stem index to identify potential LUSC-specific survival prognostic biomarkers and therapeutic targets.

2 Methods

2.1 Data processing

We first retrieved LUSC (n=502) and LUAD (n=533) RNA sequencing datasets, and the clinical information of corresponding LUSC patients and 59 healthy volunteers from the TCGA database (<https://portal.gdc.cancer.gov/>).

2.2 Differentially expressed genes

The “limma” package was selected for DEGs analysis in TCGA-LUSC and TCGA-LUAD cohorts. For processing, a $|\log_2(\text{fold change})| > 0.5$ and adjusted P-value < 0.05 were considered the cut-off criteria for screening the DEGs between the tumor and normal samples. The “heatmap” package of the R program was used to

generate a heatmap of the top 100 DEGs. Additionally, we employed a Venn diagram to indicate the specific DEGs in LUSC.

2.3 Functional enrichment analyses

The Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses of LUSC-specific DEGs were conducted with the “clusterProfiler” package of the R software; $P < 0.05$ was the statistical significance threshold. The bubble plot, circle graph, and heat map were plotted using R to visualize the top enrichment GO terms and KEGG networks. To explore the pathways and GO functions of unique differential genes in LUSC, the R “clusterProfiler” package was used for enrichment analysis based on KEGG and GO to search for common functions among DEGs, as well as related pathways of several genes. Statistical methods were used to calculate the cumulative hypergeometric distribution to analyze, within a group of genes, whether overrepresentation occurs on a functional node, as follows:

$$P(X > q) = 1 - \sum_{x=1}^q \frac{\binom{n}{x} \binom{N-n}{M-x}}{\binom{N}{M}}$$

where ‘N’ is the total gene number within the annotation system, ‘n’ is the gene number annotated by the node or pathway itself to be examined, ‘M’ is the size of the DEGs set, and ‘x’ is the number of intersections between gene sets and nodes or pathways.

2.4 Protein–protein interaction axis for LUSC DEGs

The association between LUSC-specific DEGs was predicted using STRING (<https://string-db.org>). The PPI axis was visualized with the Cytoscape software at a confidence of 0.9. In the PPI network, the individual DEG’s adjacent node numbers were computed, and the top 20 DEGs were displayed using a bar plot according to the number of adjacent nodes. Weighted gene co-expression network analysis (WGCNA) was conducted to screen out relevant modules and hub genes, which were used to develop the prognostic signature. TCGA and GTEx data based gene expression profiling interactive analysis (GEPiA) was used to predict gene interactive and customizable functions.

2.5 Construction and validation of a gene signature constructed from LUSC-specific DEGs

Based on the number of connections, the top 100 LUSC-specific DEGs were selected for subsequent analyses. We extracted the expression data of the 100 LUSC-specific genes from TCGA-LUSC patients and combined them with the clinical information of corresponding patients. The corresponding patients with TCGA-LUSC were randomly divided into a training cohort (TC, n = 336) and a validation cohort (VC, n = 145) in a 7:3 ratio. To identify

prognostic genes in LUSC, we conducted a univariate Cox regression analysis on 100 LUSC-specific DEGs. Those with a $P < 0.05$ were considered correlated with the LUSC prognosis. Subsequently, we used the least absolute shrinkage and selection operator (LASSO) and Cox regression analyses to obtain the genetic profile with the most significant prognosis from the LUSC-specific DEGs within the TCGA-LUSC patient population via the “glmnet” package in R. Individual patient risk score (RS) was computed based on the levels of the prognostic signature genes and the associated coefficients obtained from the LASSO-Cox regression model. LUSC patients were categorized into high risk (HR) and low risk (LR) cohorts based on the median RS. The overall survival (OS) of the different risk cohorts was analyzed using Kaplan–Meier analysis with the log-rank test using the “Kaplan–Meier survival” package in R. Moreover, the time dependent receiver operating characteristic (ROC) curve was generated via the “survival ROC” package in R to demonstrate the effectiveness of the genetic profile. To analyze the relationship between predictive and response variables, we employed the uni- and multivariate Cox regression analyses.

2.6 Single sample gene set enrichment analysis

The relative tumor infiltration levels of 29 immune-linked gene sets (16 immune cell types and 13 immune-linked pathways) between HR and LR groups were quantified by ssGSEA. The analysis was conducted using the “gsva” R package. Comparisons between the HR and LR cohorts were carried out via the Wilcoxon test.

2.7 Tumor stem cell index analysis

The mRNA expression based stemness index (mRNasi) and epigenetically regulated mRNasi (EREG-mRNasi) in LUSC samples were computed using the OCLR algorithm for research on gastric cancer (13) and NSCLC (14). Subsequently, the differences in mRNasi and EREG-mRNasi between the HR and LR cohorts were compared using the Wilcoxon test. The two independent stemness indices range from 0 to 1, with a value closer to 1 suggesting stronger characteristics of CSCs.

2.8 Generation of the ceRNA axis

Differentially expressed lncRNAs (DE-lncRNAs) between tumor and healthy samples were recognized as follows: $|\log_2(\text{fold change})| > 1$ and $P\text{-value} < 0.05$. The target miRNAs of lncRNAs were estimated via the miRcode database (<http://www.mircode.org/>), and the target miRNAs of prognostic DEGs were estimated via the miRanda database (<http://www.microrna.org/microrna/home.do>). The common miRNAs predicted by the miRcode and miRanda databases, as well as the corresponding lncRNA and prognostic DEGs, were input into Cytoscape software to construct a ceRNA network.

2.9 Cell culture

LTEP-s, BEAS-2B, and NCI-H520 cells were purchased from the American Type Culture Collection (ATCC) and cultured in DMEM (HyClone, USA). NCI-H520 cells were cultured in RPMI-1640 medium (Biological Industries, USA). All cells were supplemented with 10% fetal bovine serum (FBS; Biological Industries, USA) and 1% penicillin/streptomycin (Sigma, USA) and cultured under standard culture conditions (37 °C, 5% CO₂) in culture medium recommended by the ATCC.

2.10 RNA extraction and real-time polymerase chain reaction assay

Total RNA was extracted using TRIzol Reagent (Invitrogen, Carlsbad, CA) according to the manufacturer’s instructions. cDNA was synthesized using random primers and the PrimeScript RT Reagent Kit (Takara, China). Real-time polymerase chain reaction (qPCR) was performed using SYBR Premix Ex Tag (Takara, China). The PCR conditions were as follows: 95 °C for 15 s followed by 40 cycles of 95 °C for 5 s and 60 °C for 30 s. β -actin was used as the internal control. The primer sequences for real-time PCR are listed in Table 1.

2.11 Cell transfection

Small-interfering RNA (siRNA) oligonucleotides for FGG were designed and synthesized by Jima Bio (Suzhou, China). The primer sequences for the siRNAs are listed in Table 2. Transient transfection was performed using Lipofectamine 2000 Reagent (Invitrogen, USA) according to the manufacturer’s instructions. After transfection for

TABLE 1 Primer sequences for real-time PCR used in the study.

Primer name	Primer sequences (5´–3´)	
FGG	Forward Primer	TTATTGTCCAACCTACCTGTGGC
	Reverse Primer	GACTTCAAAGTAGCAGCGTCTAT
FGA	Forward Primer	AGACATCAATCTGCCTGCAAA
	Reverse Primer	AGTGGTCAACGAATGAGAATCC
JUN	Forward Primer	TCCAAGTGCCGAAAAAGGAAG
	Reverse Primer	CGAGTTCTGAGCTTCAAGGT
CPSF4	Forward Primer	CATCGGGGTCATGCAGAGTC
	Reverse Primer	CTCGCCACACTTGTAACAGGT
HIST1H2BH-1F	Forward Primer	TCACCTCCAGGAGATCCAG
	Reverse Primer	TTTGGGTTTGAACATGCGTCC
C3	Forward Primer	GGGGAGTCCCATGTACTCTATC
	Reverse Primer	GGAAGTCGTGGACAGTAACAG
CST3	Forward Primer	GTCGGCGAGTACAACAAAGC
	Reverse Primer	CACCCAGCTACGATCTGC

TABLE 2 Primer sequences for siRNA used in the study.

Primer name	Primer sequences (5'–3')	
FGG-homo-935	sense	CCUACUGGCACAACAGAAUTT
	antisense	AUUCUGUUGUGCCAGUAGGTT
FGG-homo-768	sense	GCGGGCUUUACUUUAUUAATT
	antisense	UUAUUAAGUAAAGCCCGCTT
FGG-homo-1361	sense	GGUUAUGUAAUGGCAUUATT
	antisense	UAAUGCCAUAUCAUAACCTT

48 h, cells were used for functional assays, including migration, invasion, RNA extraction, and Western blotting.

2.12 Cell proliferation assay

Cells were seeded in 96-well plates at 1×10^3 cells per well and cultured in a final volume of 100 μ L of culture medium supplemented with 10% FBS. The cell proliferation was determined using CCK-8. After incubation for 24, 48, 72, and 96 h, 20 μ L of CCK-8 reagent was added for 3 h, and the absorption at a wavelength of 490 nm was determined.

2.13 Cell cycle assay

The cell suspension was diluted to 5×10^6 cells/mL, the supernatant was removed, and 70% 500 μ L of cold ethanol was added and placed in a refrigerator at 4°C for 2 h. The cell pellet was mixed with 100 μ L RNaseA (Solarbio, China) and placed in a 37°C water bath for 30 min; PI staining buffer was added in the dark for 30 min at 4°C. Red fluorescence at 488 nm was detected using flow cytometer.

2.14 Cell apoptosis assay

The cell culture medium was collected into a centrifuge tube. The cells were digested with Edta-free pancreatic enzymes and added into the cell culture medium, centrifuged, and precipitated. The cells were then re-suspended with 1 mL PBS precooled at 4 °C and the precipitated cells were centrifuged again. The cells were re-suspended with 1x binding buffer and the concentration was adjusted to 5×10^6 /mL; 100 μ L cell suspension was added to a 5 mL flow tube, mixed with 5 μ L Annexin V/FITC (Solarbio, China), and incubated at room temperature for 5 min in the dark. A total of 5 μ L propyl iodide solution (PI) and 400 μ L PBS were added for immediate flow detection.

2.15 Wound healing assay

Cells were placed in 12-well plates. When cells grew to 90–95% confluence, cell monolayers were wounded by scratching with

plastic micropipette tips and washed twice with PBS. The cells were rinsed with PBS and cultured in DMEM or RPMI 1640 supplemented with 1% FBS. Images of the different stages of wound healing were obtained via microscopy at 0, 24, and 48 h. Relative cell motility was quantified using Image-Pro Plus.

2.16 Transwell migration and invasion assay

Cell migration and invasion assays were performed in 24-well plates with 8- μ m-pore size chamber inserts (Corning, USA). For the migration assays, 5×10^4 cells in 200 μ L of serum-free culture medium were seeded into each well of the upper chamber with the noncoated membrane, and 800 μ L of medium supplemented with 10% FBS was added to the lower chamber. For invasion assays, 1×10^5 cells in 200 μ L of serum-free culture medium were seeded into each well of the upper chamber with the Matrigel-coated membrane, while 800 μ L of medium supplemented with 10% FBS was added to the lower chamber. Cells that migrated through the membrane were fixed with 100% methanol, stained with 0.1% crystal violet for 30 min, imaged, and counted under a light microscope (Olympus, Japan).

2.17 Western blot assay

Cells grown in 6-well plates were lysed on ice using RIPA buffer. The lysis mixtures were centrifuged, and the supernatants were collected. Total protein was separated using SDS-polyacrylamide gel electrophoresis and transferred onto PVDF membranes (Millipore, USA). After blocking the membranes with non-fat milk, the membranes were incubated overnight with the following primary antibodies: anti-N-cadherin (1:1,000), anti-E-cadherin (1:1,000), anti-GAPDH (1:1,000) (Abcam, UK). The membranes were then incubated with horseradish peroxidase-conjugated secondary antibodies (1:2,000). The analysis was performed using an enhanced chemiluminescence system (Bio-Rad, USA). Binding was analyzed using Image J 6.0.

2.18 THP-1 cell infiltration

THP-1 cells were seeded at 1×10^6 per well in 6-well plates and treated with PMA (100 nmol; Sigma-Aldrich, USA) for 48 h. M1 macrophages were polarized by incubation with INF- γ (20 ng/mL; R&D System, USA) and LPS (100 ng/mL; Sigma, USA) for 48 h.

After transfection with si-NC or si-FGG in the absence or presence of coculture, a cell migration assay was conducted using 24-well Transwell plates (8.0 μ m; Corning, NY, USA). The macrophages or cancer cells (5×10^4 , LTEPs-si-NC, LTEPs-si-FGG) were planted into the upper chambers, while 600 μ L RPMI 1640 containing 10% FBS were placed into the lower chambers. Thereafter, the Transwell plates were incubated at 37 °C, 5% CO₂ for 48 h, fixed in 4% formaldehyde for 30 min, and stained with 0.01% crystal violet. Non-migrating cells were carefully removed

with a cotton swab, while cells that migrated to the lower chambers were counted under a microscope.

2.19 Statistical analysis

All data analyses were conducted using the R language (version 3.5.1). The levels of immune checkpoint genes between the HR and LR cohorts were compared using the Wilcoxon test. Uni- and multivariate analyses were employed to screen for stand-alone prognostic markers for LUSC survival. $P < 0.05$ was set as the significance threshold.

3 Results

3.1 Identification of specific DEGs for LUSC

We analyzed DEGs between tumor and normal samples in TCGA-LUSC and TCGA-LUAD cohorts. Overall, 2,878 DEGs (1,466 upregulated and 1,412 downregulated) were identified in LUSC, relative to normal samples (Figure 1A). In addition, 1,629 DEGs were identified in LUAD, among which, 714 were highly expressed and 915 were scarcely expressed (Figure 1B). The top 100 DEGs in LUSC (Figure 1C) and LUAD (Figure 1D) are shown in the heat maps. We further applied an online Venn diagram to identify LUSC-specific DEGs (Figure 1E). Consequently, 1,604 specific DEGs for LUSC were obtained, as shown in a heat map (Figure 1F).

3.2 FEA and PPI analysis of the novel biomarkers in LUSC

To elucidate the physiological activities of these LUSC-specific DEGs, GO and KEGG enrichment analyses were carried out. GO terms revealed that these LUSC-specific DEGs were markedly enriched in immune-linked biological systems such as T cell-mediated immunity, immune response-related neutrophil activation, neutrophil degranulation, neutrophil-based immunity, and neutrophil activation (Figure 2A). KEGG analysis revealed that the LUSC-specific DEGs were associated with melanogenesis, small-cell lung cancer, the PI3K-Akt axis, viral myocarditis, human papillomavirus infection, ECM-receptor association, the Rap1 signaling pathway, Staphylococcus aureus infection, and glutathione metabolism (Figure 2B). PPI interaction networks containing 1,604 nodes and 14,209 edges further revealed the interactions between these LUSC-specific DEGs (Figure 2C). The top 20 DEGs are displayed in a bar plot based on the quantity of adjacent nodes (Figure 2D). The top 100 genes of connectedness were obtained using a collateral analysis. The genes with the top 100 connectedness were single factors. Then, Cox and LASSO regression analyses were employed for risk model construction. WGCNA was used to analyze the hub genes' biological behavior,

and the correlation between alteration in hub gene expression and clinical characteristics was confirmed via external data from the GEPIA Database (<http://gepia.cancer-pku.cn/>). The results of WGCNA and GEPIA for hub genes suggested that all hub genes were significantly elevated in tumor tissues. Following the adjustment of confounding factors, we developed a prognostic profile using three genes with remarkable predictive ability.

3.3 Prognostic signature of LUSC-specific DEGs

Based on the counts of connections, the top 100 LUSC-specific DEGs in the PPI network were selected for further analysis. To identify the prognostic genes in LUSC, we further employed a univariate analysis of the 100 LUSC-specific DEGs. Eight were associated with the prognosis of patients with LUSC ($P < 0.05$); univariate Cox regression analysis results are shown in Supplementary Table 1. FGG, C3, FGA, ORM1, JUN, and CST3 served as risk hazards ($HR > 1$), whereas CPSF4 and HIST1H2BH served as a protective function ($HR < 1$) in LUSC (Figure 3A). LASSO analysis was employed to improve the robustness of the eight LUSC-specific DEGs. Eight genes carrying a P -value < 0.05 in the univariate Cox analysis were used to construct a LASSO regression. To reduce the feature dimension, we used the R software's "glmnet" package, set the parameter family as Cox, realized LASSO logistic regression, selected strong correlation features, and obtained the two graphs depicted in Figure 3; one is the graph of gene coefficient, and the other is the error graph of cross-validation. As shown in Figure 3B, the seven characteristic genes with a lambda.min of 0.0134 were FGG, C3, FGA, JUN, CST3, CPSF4, and HIST1H2BH. The seven genes and their corresponding coefficients were selected as the most prognostic gene signatures in LUSC. We further calculated the RS for individual patients with LUSC using the expression of the seven prognostic genes and associated coefficients retrieved from the LASSO-Cox analysis; LASSO analysis was then employed for characteristic genes and coefficients screening, as shown in Supplementary Table 2. Subsequently, the median of the RSs was utilized as a standard to separate the LUSC patients into HR and LR cohorts in both the TC and VC. The risk curve and distribution of OS status are shown in Figure 3C. Moreover, the expression patterns of the seven prognostic genes in the HR and LR cohorts verified the prognostic value of the seven markers. Figure 3C consists of three parts: upper (a), middle (b), and lower (b), all of which demonstrate that the HR cohort exhibited an elevated survival RS. Notably, the Kaplan–Meier analysis indicated that LR LUSC patients exhibited a markedly higher survival probability, compared to the HR cohort (Figure 3D; $P < 0.05$). The results of ROC analysis further tested the TC, which showed that this genetic profile could effectively estimate the OS of LUSC (Figure 3E).

The VC was also tested, and the risk curve and distribution of OS status are shown in Supplementary Figure 1. The survival and ROC curves of VC are shown in Figures 3F, G.

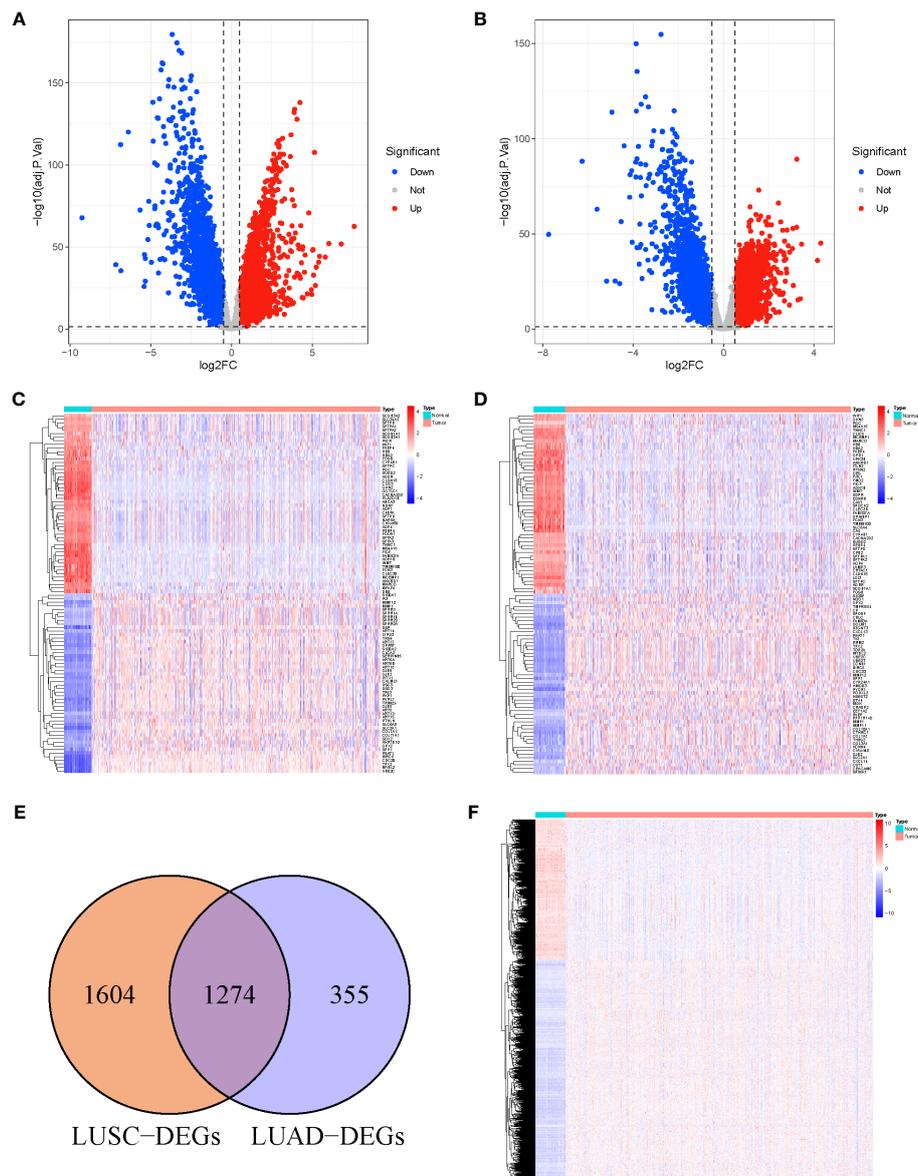


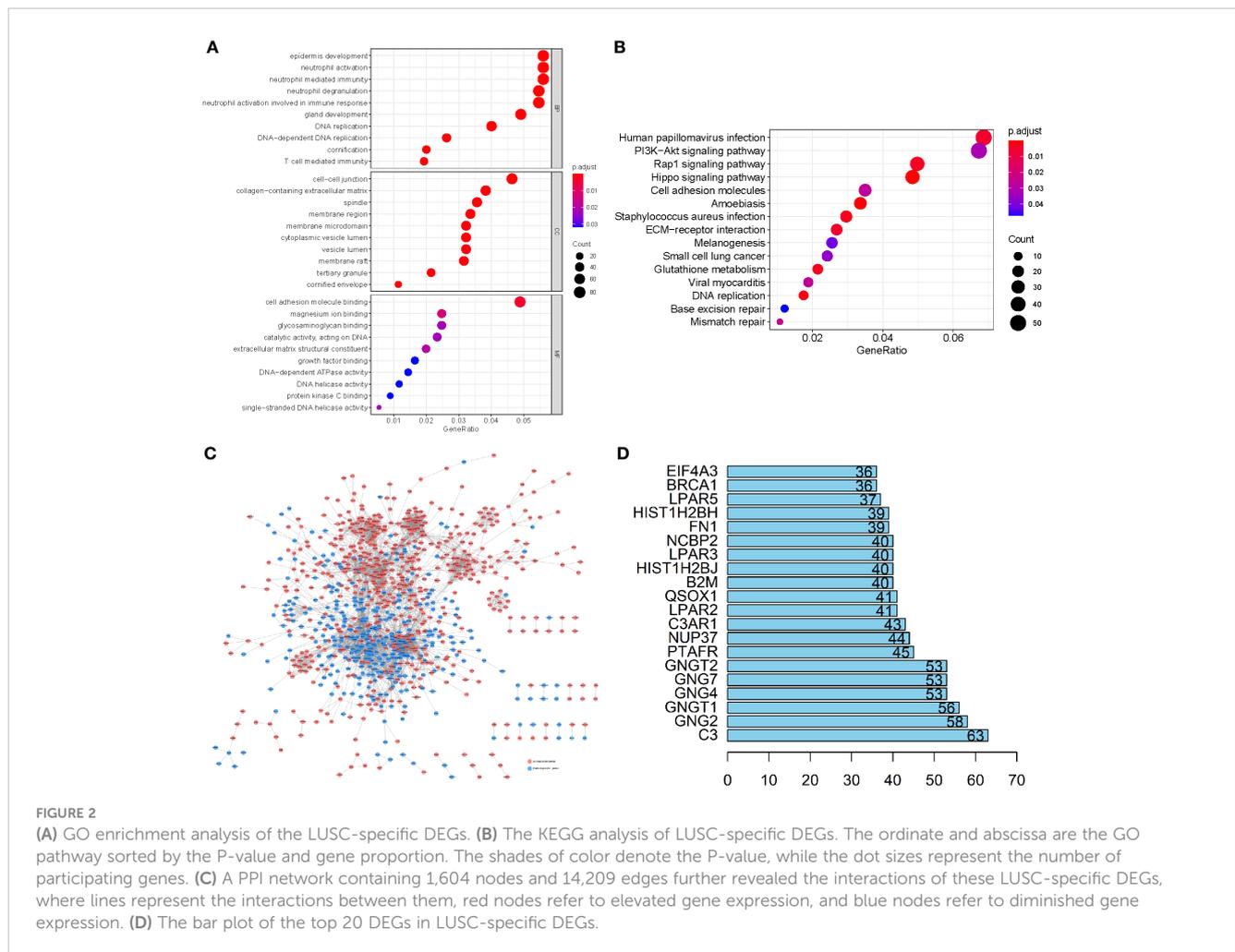
FIGURE 1

(A) Volcano Plots of 2,878 LUSC-DEGs. (B) Volcano Plots of 1,629 LUAD-DEGs. Multiples of the abscissa difference (Tumor/Normal) taken the logarithm of 2 and the ordinate representation of $-\log_{10}(\text{adj.P.Val})$. Each dot represents a gene. Red dots indicate gene upregulation (Tumor vs. Normal samples), blue dots indicate downregulation (Tumor vs. Normal samples), and gray dots indicate no significant differences in expression. (C) The heat map of the top 100 DEGs in LUSC. (D) The top 100 DEGs LUAD. The abscissa direction represents the DEGs, while the vertical direction represents the samples. Colors indicate normalized differential expression; high and low expressions are shown in red and blue, respectively. (E) The Venn diagram of 1,604 LUSC-specific DEGs calculated by subtraction of LUSC-DEGs and the cross-section of LUSC and LUAD DEGs. (F) The heat map of LUSC-specific DEGs. The abscissa direction indicates the DEGs, while the vertical direction indicates the samples. Colors indicate normalized differential expression; red represents elevated levels, and blue represents reduced levels.

3.4 The seven-gene signature of LUSC represents an independent stand-alone prognostic value

To elucidate whether the prognostic gene profile was independent of clinicopathological features such as age, pathological stage, and TNM stage, univariate and multivariate analyses were conducted. Univariate analysis revealed that age, as well as pathologic, pathologic T, and pathologic M stages, were strongly correlated with LUSC patients' OS (Figure 4A).

Multivariate analysis based on the above clinicopathological characteristics further revealed that the RS was directly correlated with OS (Figure 4B; $P < 0.001$). The predictive efficiency of these clinicopathological characteristics was evaluated using ROC analysis, and the RS was employed as a predictor stand-alone indicator of LUSC outcome (Figure 4C). We observed marked differences between T1 and T2 of the T stage, and N0 and N1 of the N stage (Supplementary Figures 2, 3); however, there was no significant difference in other periods (Supplementary Figures 4–6). To explore the independent prognosis of risk models and



clinicopathological factors, the Cox-independent prognostic analysis of age, the T, M, and N staging, and the RS showed that Pathologic_M and RiskScore were stand-alone prognostic indicators for LUAD ($P < 0.05$). We next analyzed RSs and various clinical features, including age, sex, tumor stage, T (size or extent of the tumor itself), M (distant metastasis), and N (tumor peripheral lymph node invasion and metastasis). A differential expression heat map of the genes was drawn (Figure 4D). Univariate and multivariate results were consistent, indicating that the conclusions were stable and easy to interpret.

3.5 Characteristics of immune infiltration in LUSC

Previous research has revealed a relationship between immune cell invasion and clinical prognosis in cancers, which may be utilized as drug targets to enhance the prognosis of patients (15, 16). Therefore, we quantified the tumor infiltration levels of 29 immune-related gene sets in the HR and LR cohorts. Immune checkpoint inhibitors were reported to be effective potent therapeutic methods against various cancers (17–19); hence, we assessed the levels of key immune checkpoint molecules in LUSC.

The HR cohort was markedly correlated with elevated tumor infiltration levels in LUSC (Figures 5A, B; all P-values < 0.05); however, the tumor infiltration levels of NK cells showed no significant differences between the HR and LR cohorts. Importantly, the checkpoint scores between the HR and LR cohorts were significantly different. ssGSEA was performed on the samples from the HR and LR cohorts; we observed marked differences in the levels of certain immune cell infiltrates between the HR and LR cohorts. The infiltrating cells included aDCs, B cells, CD8⁺ T cells, DCs, iDCs, macrophages, mast cells, neutrophils, pDCs, T helper cells, Tfh, Th1 cells, Th2 cells, TIL, and Tregs. There were also significant differences in the levels of some immune-linked pathways between the HR and LR cohorts, such as APC_co_inhibition, APC_co_stimulation, CCR, Check-point, Cytolytic_activity, HLA, Inflammation-promoting, MHC_class_I, Parainflammation, T_cell_co-inhibition, T_cell_co-stimulation, Type_I_IFN_Reponse, and Type_II_IFN_Reponse. Immune checkpoints refer to those that inhibit cytotoxic T lymphocyte activation, or cytotoxicity, as well as T lymphocyte (killer T cell) interaction. These findings suggest that the prognostic model is related to the function of antigen-presenting cells (APCs), cytotoxic T cells, immune checkpoints, and major histocompatibility complex (MHC). Thus, the risk model could also be an indicator of tumor immune response in LUSC.

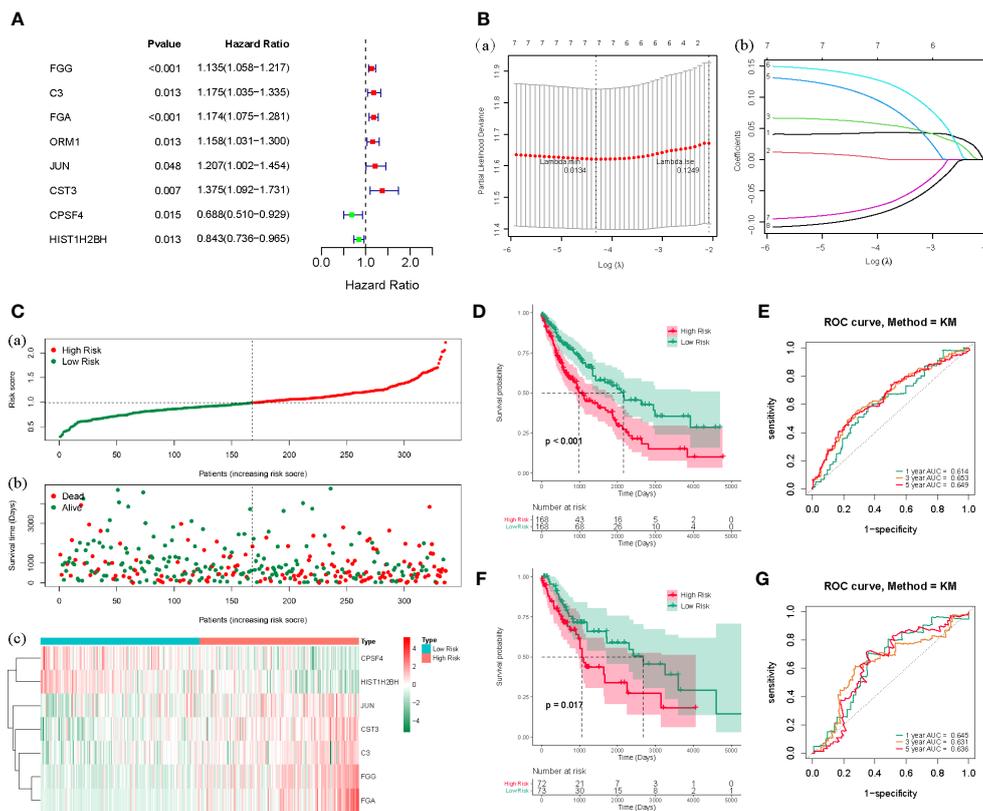


FIGURE 3

(A) The forest map of the eight risk genes (FGG, C3, FGA, CRM1, JUN, CST3, CPSF4, and HIST1H2BH) using univariate analysis. (B) (a) LASSO analysis, where the screened characteristic gene ordinate is the gene coefficient; (b) the abscissa is the log(Lambda), and the ordinate denotes cross-validation error. In the analysis, we identified the position with the minimum error of cross-validation. In (B), the dotted line on the left represents the position with the minimum error of cross-validation. Based on the position (lambda.min), we determined the associated horizontal coordinate log (Lambda) and the number of characteristic genes (shown above); we also found the optimal log(Lambda) value and the associated gene and its coefficient in the left figure (A). (C) The risk curve and the distributions of OS status of the seven-gene TC ($P < 0.05$). The risk score (RS) in high- (HR) and low-risk (LR) cohorts (a), the OS status (b), and the heat map (c) are shown. The figure above (a) is consistent with the abscissa of the middle figure (b), indicating that RSs rose from left to right. The ordinate represents the RS and survival time, while the dotted line represents the median RS and the corresponding number of patients. Below (c) is the gene expression heat maps in the HR and LR cohorts. (D) The OS curve based on the HR and LR cohorts. (E) ROC curve of the seven-gene set in TCGA-LUSC training cohort (TC) 1-3-5-years OS. (F) TCGA-LUSC validation of survival curves for concentrated HR and LR cohorts. (G) ROC curves for 1-3-5-years OS in TCGA-LUSC validation cohort (VC).

Immune checkpoint molecules for immune function are crucial for TME and immunotherapy (20). To examine the potential association between molecular levels and immune checkpoints, we analyzed the expression of several key immune checkpoint sites, including TNFRSF18, TNFRSF14, CD160, CD48, CD244, TNFSF18, TNFSF4, CD28, ICOS, PD-1 (PDCD1), CD47, BTLA, TIGIT, CD80, CD86, TIM-3 (HAVCR2), PD-L1 (CD274), CD27, LAG3, CD276, LGALS9, CD226, CD70, TNFSF14, CEACAM1, PVR, and CD40. As shown in Figure 5C, apart from TNFSF18, TNFSF4, CD274, LAG3, and CD276, the levels of most immune checkpoint genes were markedly different between the HR and LR cohorts (Figure 5; all P-values < 0.05).

3.6 Cancer stem cell characteristics of the risk model

Cancer stem cells serve essential functions in tumor survival, metastasis, proliferation, and recurrence, owing to their self-renewal

ability and production of heterogeneous tumor cells (21). MRNAsi reflects the gene expression characteristics of stem cells. We used mRNAsi as the stemness index to investigate the similarities between cancer and stem cells. The index ranged from 0 to 1; the value of mRNAsi close to 1 indicated enhanced stem cell features of the tumor cells. Thus, the mRNAsi and EREG-mRNAsi of LUSC samples were further computed using the OCLR algorithm and then compared between the HR and LR cohorts. Figures 6A, B shows significant differences in the mRNAsi and EREG-mRNAsi between the two cohorts ($P < 0.05$).

3.7 Establishment of a ceRNA network for LUSC

LncRNAs and circRNAs are generally perceived as competing endogenous RNAs (ceRNAs) that bind to miRNAs. ceRNA analysis refers to the analysis of the entire ceRNA regulatory network; usually circRNA-miRNA-mRNA analysis or lncRNA-miRNA-

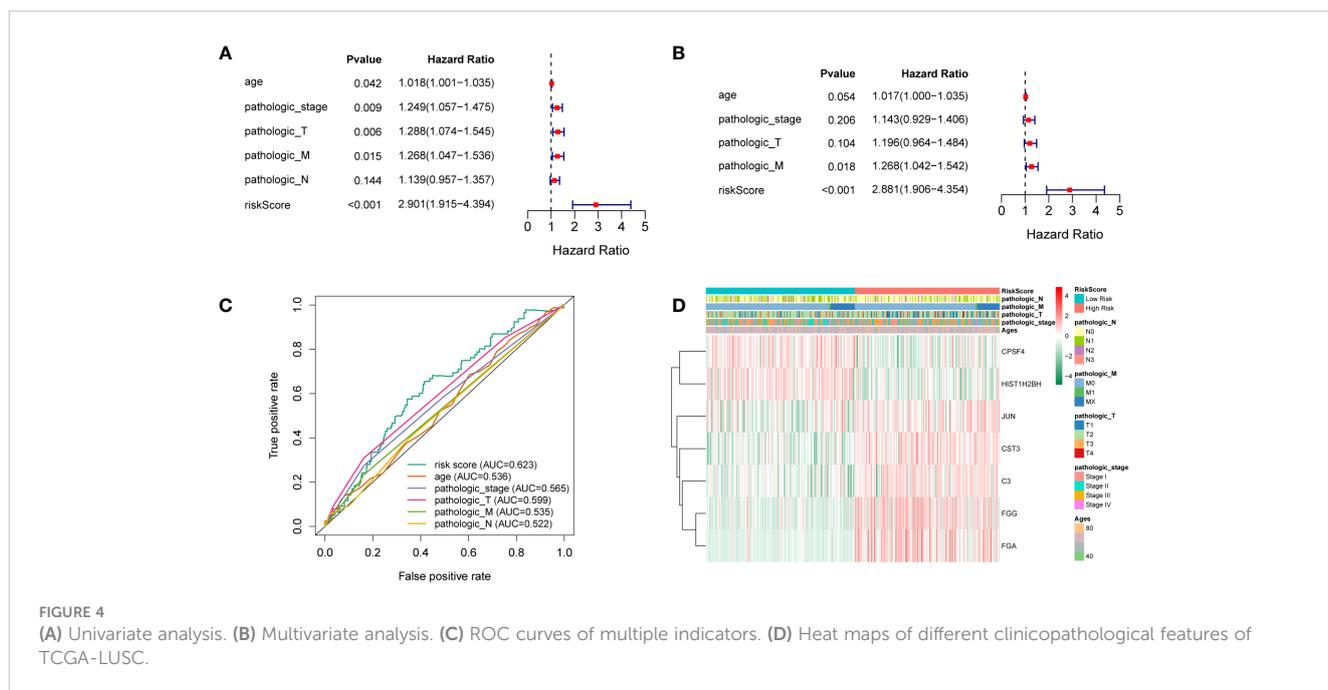


FIGURE 4 (A) Univariate analysis. (B) Multivariate analysis. (C) ROC curves of multiple indicators. (D) Heat maps of different clinicopathological features of TCGA-LUSC.

mRNA analysis is perceived as the core of the ceRNA regulatory network. With competitive binding of ceRNAs, such as lncRNA or circRNA with miRNA, the transcription level of the genes regulated by miRNAs will increase. To further elucidate the potential regulatory mechanism of these seven prognostic DEGs in LUSC prognosis, we generated a ceRNA network using the DE-lncRNAs and prognostic DEGs. The target miRNAs of DE-lncRNAs were predicted using the miRcode database, and the target miRNAs of prognostic DEGs were predicted using the miRanda database. A ceRNA network containing 19 lncRNAs, 50 miRNAs, and 7 prognostic DEGs demonstrated the molecular mechanism of LUSC-specific DEGs in LUSC prognosis (Figure 7).

3.8 FGG and Clinical Parameters in patients with LUSC

The prognostic values of FGG, C3, FGA, JUN, CST3, CPSF4, and HIST1H2BH7 genes in LUSC in the TCGA database suggest that they may play a role as key risk factors in tumors (Supplementary Figure 7). The expressions of seven prognostic genes in human LUSC cell lines NCI-H520 and LTEP-s were detected using q-PCR; FGG was significantly highly expressed in both LUSC cell lines (Supplementary Figure 8). We also examined the expression of FGG in surgically collected, paired, LUSC samples, and adjacent normal tissues from 6 patients. Remarkably, all LUSC specimens had markedly increased FGG protein levels compared with matched adjacent normal tissues (Figure 8). Our clinical observations reveal that FGG is significantly hyper-expressive in LUSC patient samples, further demonstrating the clinical value of FGG in LUSC.

3.9 FGG correlates with tumor progression, immune infiltration, and stem index in LUSC

Immunofluorescence showed that FGG was expressed in the nucleus of LUSC (Figures 9A, B). To demonstrate the biological function of FGG in LUSC cells, NCI-H520 (Figures 9C, E) and LTEP-s cell lines (Figures 9D, G) with FGG knockdown were successfully constructed.

Our results suggest that FGG can affect the tumor process of LUSC cells, as shown by the proliferation (Figures 9F, H), cloning (Figures 9I, J), invasion (Figures 9K, L), and migration (Figures 9M, N) of NCI-H520 and LTEP-s being significantly inhibited following FGG knockdown. In addition, the result of western blot showed that the expression of E-cadherin was increased while that of N-cadherin and VIMENTIN were decreased following FGG knockdown, which also corresponded to the inhibition of migration and invasion (Figures 9O, P). Subsequently, we evaluated the scores of 22 kinds of tumor immune cell infiltration in LUSC patients according to the expression of FGG and found that FGG was significantly correlated with 10 kinds of immune cell infiltration, including M1 macrophages (Figures 10A, B). *In vitro* Transwell experiments showed that the ability of NCI-H520 (Figure 10C) and LTEP-s cell lines (Figure 10D) with low FGG to recruit M1 mononuclear macrophages was significantly down-regulated. After FGG knockdown, KLF4, Nanog, CD44, and SOX2 in NCI-H520 cells were significantly decreased, while CD133 showed no significant changes (Figures 10E, F). After FGG knockdown, KLF4, Nanog, CD44, and SOX2 in NCI-H520 cells were significantly decreased, while CD133 showed no significant changes (Figure 10E). After FGG downregulation, the expressions of CD44 and CD133 in

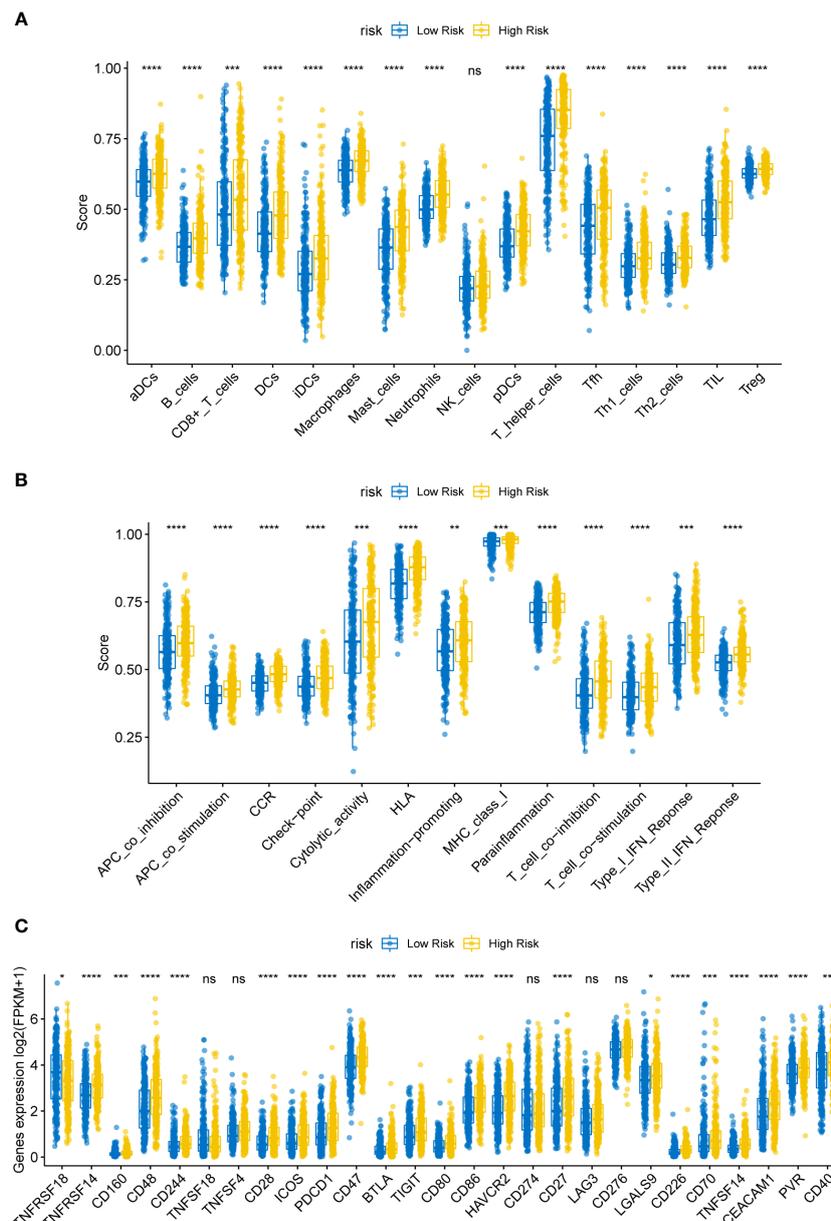


FIGURE 5
(A) Box plot of immune infiltrating cells in the high- (HR) and low-risk (LR) cohorts. The HR cohort was strongly associated with elevated tumor infiltration levels in LUSC ($P < 0.05$). **(B)** Box plot of tumor infiltrated pathway. **(C)** Box plot of immunoassay sites in the HR and LR cohorts. The levels of the remaining immune checkpoint genes were markedly different between the HR and LR cohorts (P -values < 0.05); ns, not significant. (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$).

LTEP-s cells were significantly decreased, while no significant changes in KLF4, Nanog, and SOX2 were observed (Figure 10F). These results indicate that FGG affected the tumor progression, immune infiltration, and stem index of LUSC cells.

4 Discussion

LUSC is a subtype of NSCLC and accounts for nearly 40% of all lung cancers. Early detection and the prognostic assessment of LUSC remain challenging, hence the poor 5-year survival rate of patients (22). Recent studies have improved the prognosis

prediction for LUSC patients, focusing on biomarkers. For example, Shi et al. investigated DNA methylation profiling and proposed potential diagnostic biomarkers for LUSC (23). Chen et al. investigated the roles of IRGs in the deterioration of lung cancer and indicated the distinction between LUAD and LUSC from the perspective of the immune response (24). Liao et al. identified biomarkers with cancer stem cell characteristics in LUSC (14). To date, the prognostic gene signatures for prognostic prediction of LUSC are scarce and warrant further investigations. Several studies have proposed prognostic markers for survival prediction in patients with LUSC. Zhang et al. suggested that IRGPI could be used as a prognostic marker (25), while Li et al.

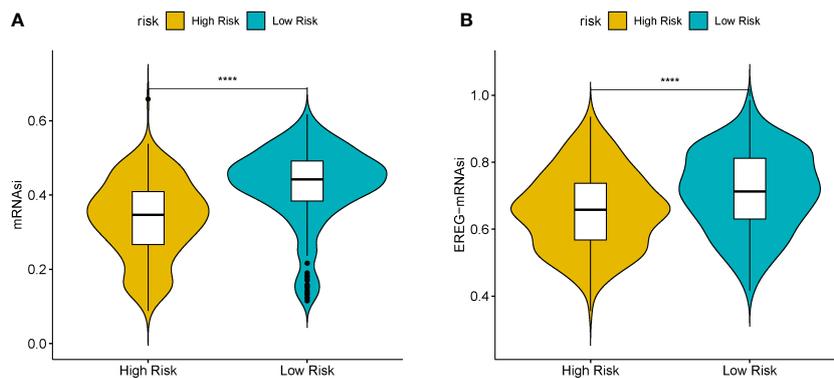


FIGURE 6 (A) Boxplots of mRNAasi in the high- (HR) and low-risk (LR) LUSC patients. (B) Boxplots of EREG-mRNAasi in the HR and LR LUSC patients. (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$, **** $P < 0.0001$).

constructed an mRNA signature to predict the outcomes of patients with LUSC (26). Liu et al. have identified an miRNA signature with potential clinical implications in the outcome prediction of LUSC (27). Indeed, several lncRNAs, such as VPS9D1-AS1 and MALAT-1, are correlated with the survival of LUSC patients (28, 29). Huang et al. reported a nine-long non-coding RNA signature for prognosis prediction of patients with LUSC (7). However, no prognostic indicators of LUSC have been established based on tumor progression, immune infiltration, and stem index analysis.

Recent studies have found that LUSC differs from LUAD in terms of genomic, epigenetic, CSC stemness, and TME characteristics. According to previous research, CSCs may lead to

cancer recurrence and drug resistance (30, 31). The TME is a mutually adaptive environment in which tumor cells escape immunological surveillance. Tumor progression involves crosstalk between CSCs and the TME (32, 33), such as the induction of CSCs in EMT (34) and the interaction of angiogenesis and components of the TME (35). Herein, we adopted a comprehensive perspective of cancer biology based on tumor progression, TME, and CSC index for a better understanding of LUSC as an independent NSCLC from different dimensions. We recognized the importance of the particularly expressed genes in LUSC based on the TCGA database and DEGs in HR and LR cohorts; from this, we recognized the functions of independent genes as potential

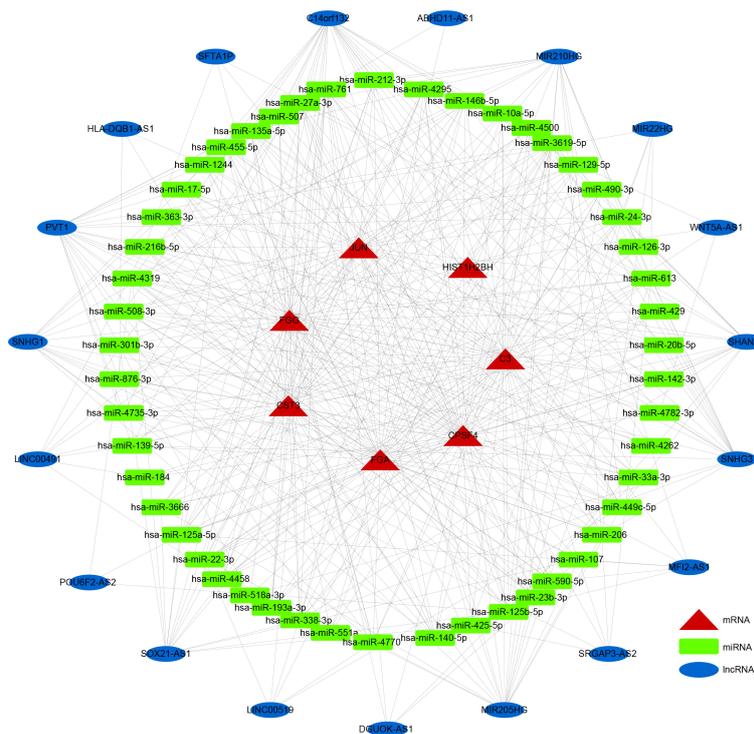


FIGURE 7 ceRNA network. Rectangles, ellipses, and triangles represent the miRNAs, lncRNAs, and mRNAs of the risk model genes, respectively.

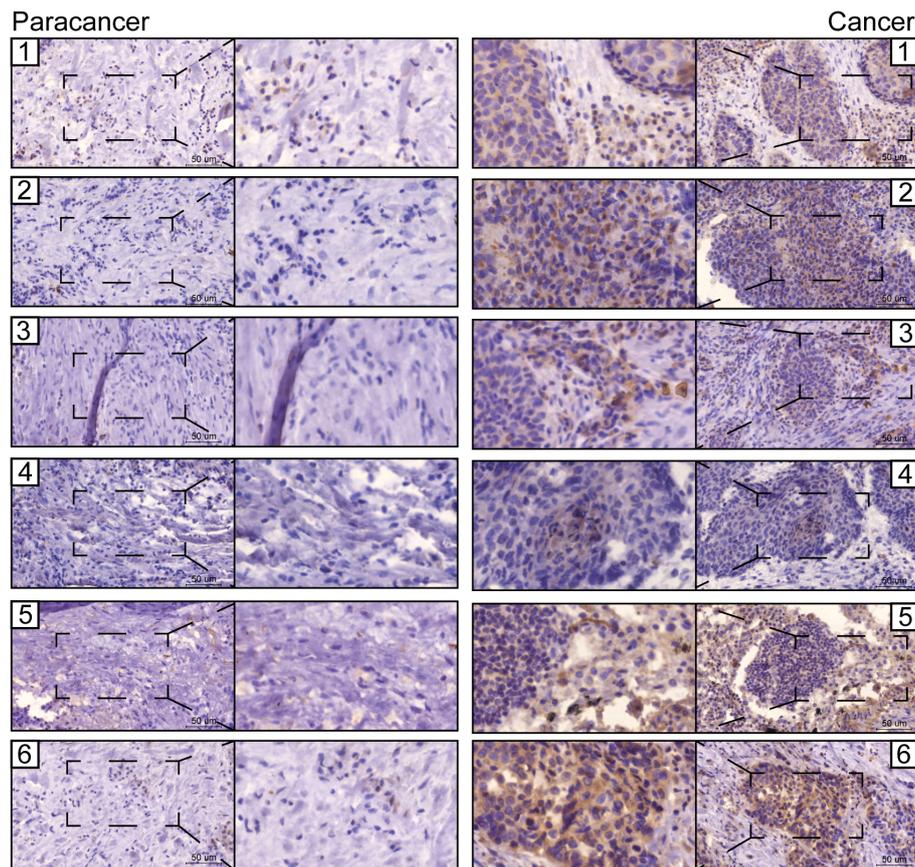


FIGURE 8
Representative images from immunohistochemical staining of FG2 in lung cancers ($n = 6$) and normal tissues ($n = 6$). Scale bars: 100 μm and 50 μm .

predictors of tumor invasion, metastasis, tumor stem cell characteristics, and immune cell infiltration. Seven prognostic genes were varied in LUSC and were associated with the TNM stage and prognosis; these genes were FG2, C3, FGA, JUN, CST3, CPSF4, and HIST1H2BH. FG2, C3, FGA, JUN, and CST3 were associated with poor outcomes in LUSC patients, whereas CPSF4 and HIST1H2BH served as positive prognostic markers in LUSC patients.

In terms of clinicopathological features, the seven-gene biomarkers showed differences in tumor metastasis and invasion, and the significant differences between T1 and T2 of the T stage and N0 and N1 of the N stage suggested that the modification occurred during the early stage of tumor disease; however, the factors of dabbling were limited, such as the lack of the status of smoking status, driver factors, ORR of the various chemotherapy, and immune checkpoint blockade subgroups. Kaplan–Meier analysis showed that LUSC patients in the LR group exhibited significantly higher OS than those in the HR group, while ROC curve analysis results showed that this gene profile could effectively predict the OS of LUSC. Subsequently, our independent prognostic value analysis showed that protective genes were highly expressed in the low-risk group, while the risk genes were highly expressed in the high-risk group, indicating stable results. Moreover, the ROC curve showed that RS could be used as an independent prognostic factor

effectively predicting LUSC outcomes. We also analyzed the relationships between HR and LR cohorts, immunoinfiltrating cells, and immune pathways, and showed that HR patients exhibited significantly elevated levels of tumor cell immune infiltration and that the molecular expression of immune checkpoint genes significantly differed between HR and LR patients.

Next, we analyzed the stem cell characteristics of the model and showed that mRNA was associated with prognosis and relevance; significant differences were noted in mRNA_i and EREG-mRNA_i between HR and LR patients, providing new insights into the clinical features, immune response, and TME of tumors based on the dry index. Finally, we constructed a ceRNA network containing 19 lncRNAs, 50 miRNAs, and 7 prognostic DEGs, demonstrating the prognostic value of novel biomarkers for Lusc-specific DEGs.

The prediction of the risk prognostic model constructed can potentially provide more reliable theoretical support for clinical application. However, bioinformatics is only a short practical perspective to this goal; therefore, we conducted specific molecular studies on prognostic genes. Based on the risk model constructed above, combined with RT-qPCR assay and survival analysis of the TCGA database, we screened LUSC-specific prognostic genes and found that FG2 was closely correlated with LUSC results in univariate Cox analysis ($P=0.000427708$), and mRNA levels of FG2 were stably expressed in NCI-H520 and

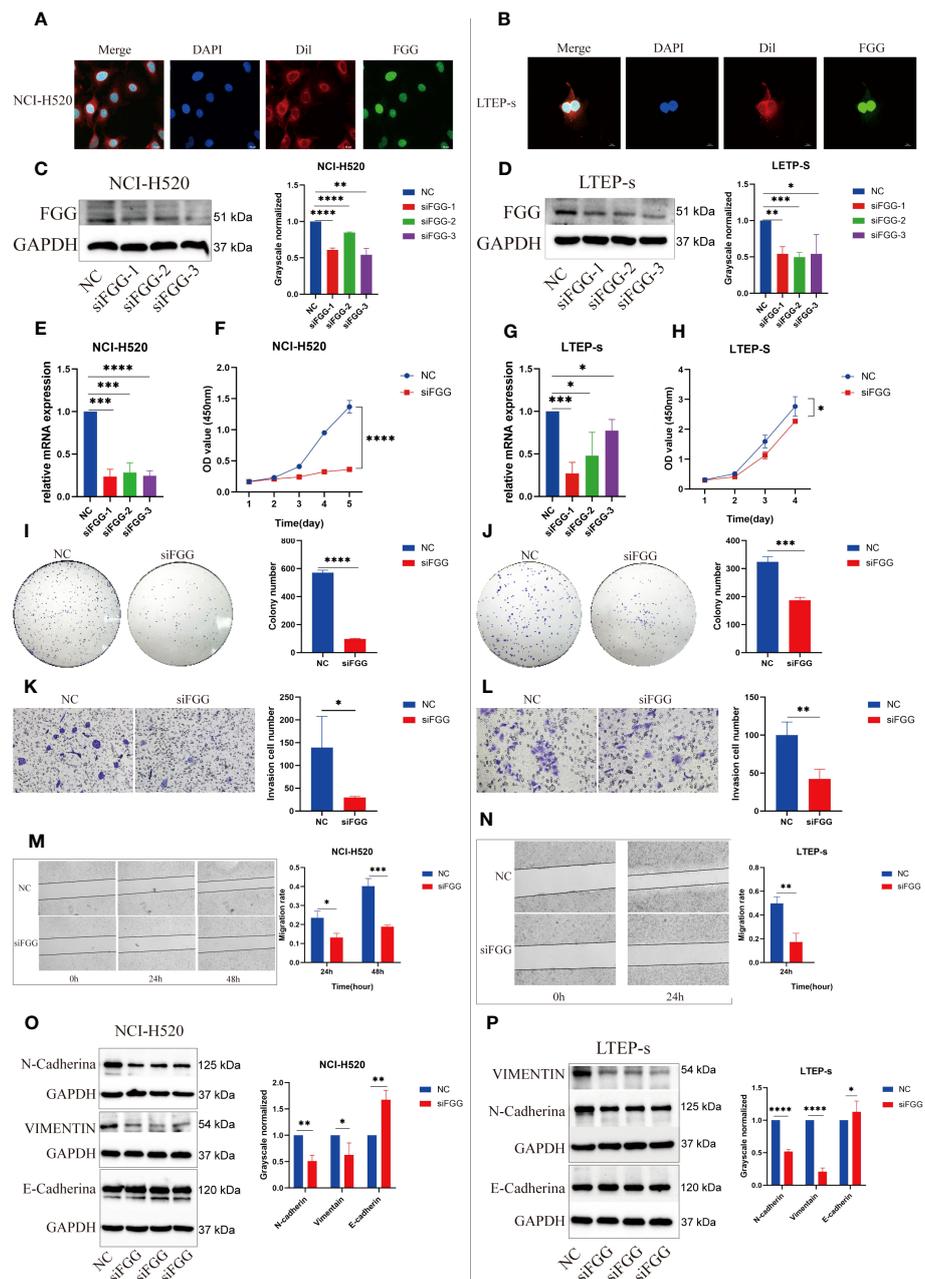


FIGURE 9 Expression locations of FG detected using immunofluorescence in (A) NCI-H520 and (B) LTEP-s. FG knockdown was determined using western blotting in (C) NCI-H520 and (D) LTEP-s cells. FG knockdown was determined using Q-PCR in (E) NCI-H520 and (G) LTEP-s cells. CCK-8 assay was used to detect the proliferation of (F) NCI-H520 and (H) LTEP-s cells. Viability line graph (I) NCI-H520 and (J) LTEP-s cell colony formation result. The result of the invasion of (K) NCI-H520 and (L) LTEP-s cells. The results of Wound Healing and migration of (M) NCI-H520 cells and (N) LTEP-s cells. Western blotting assay showing EMT markers N-cadherin, Vimentin, and E-cadherin expression following FG knockdown in (O) NCI-H520 and (P) LTEP-s cells. The significant differences were analyzed using GraphPad Prism *t*-test, *n*=3 (**P*<0.05, ***P*<0.01, ****P*<0.001, *****P*<0.0001).

LTEP-s cells and significantly up-regulated compared with normal airway epithelial cells. Therefore, their roles in tumor progression, immunoinfiltration, and dry characteristics were further analyzed.

FGG is the γ -chain of fibrinogen, a large, complex glycoprotein with a total molecular mass of approximately 340 kDa, comprising three pairs of polypeptide chains: α A (encoded by the FGA gene), β B (FGB), and γ (36). FGG has a conserved globular domain, γ C, at the COOH terminus, which is a major integrin binding site for

fibrinogen. Yokoyama et al. found that the C-terminal region of FGG, as the primary integrin binding site of fibrinogen, participated in the process of thrombosis, angiogenesis, and inflammation (37, 38). Nobuaki Akakura et al. found that isolated γ C and its mutant γ C399tr induce endothelial cell apoptosis, and recombinant soluble γ C399tr inhibited tumor growth, intratumoral vascular development, and metastasis *in vivo* (39). Previous studies have shown that fibrinogenemia, as a prognostic factor (40–42), is often

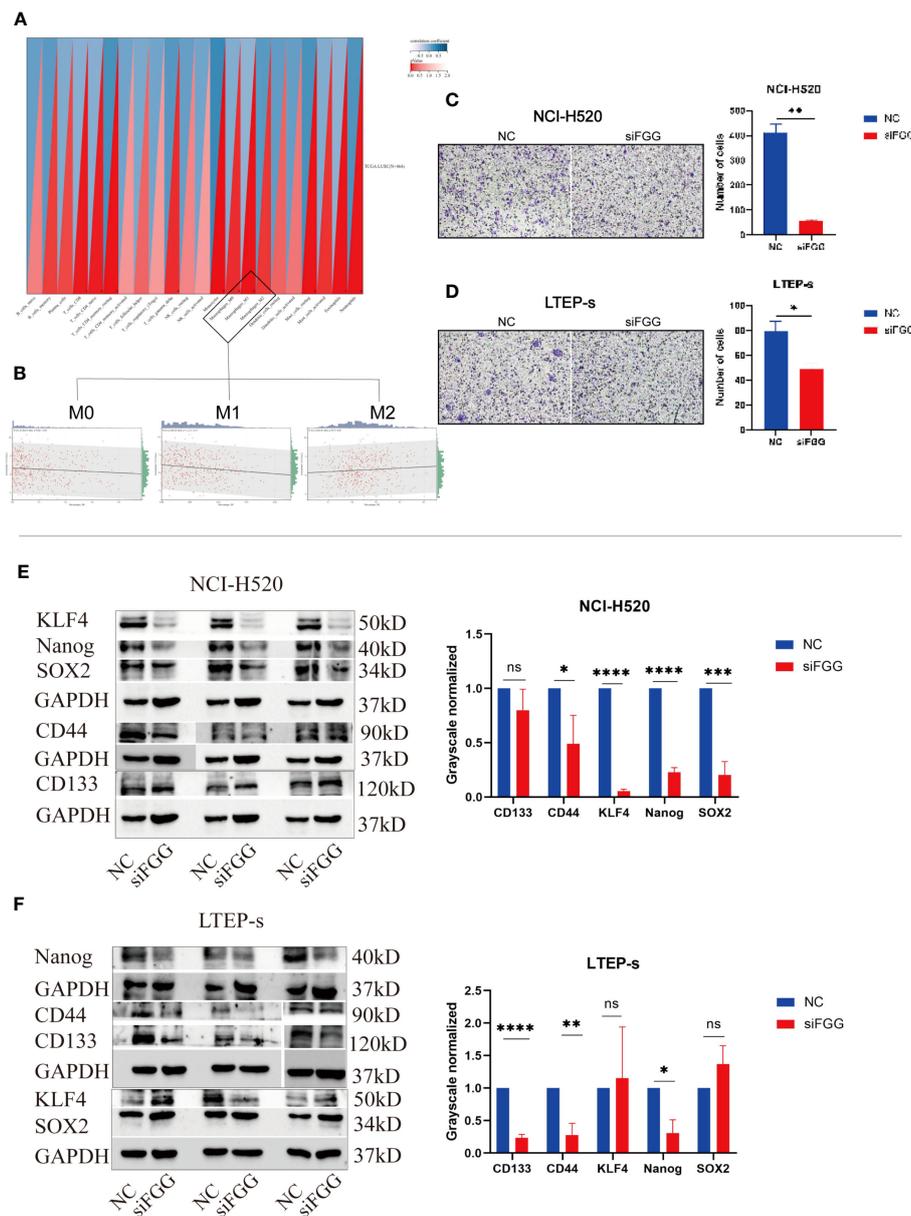


FIGURE 10
(A) Pearson's correlation coefficient of FG G with 22 immune cell infiltration scores in LUSC was calculated using the corr.test function of the R package psych (version 2.1.6), and 10 significantly correlated immune infiltration scores, including macrophages, were identified, **(B)** for further individual correlations plotted for FG G with M0, M1, and M2 macrophages, respectively ($p=0.04$, $r=0.09$, $p=1.1e-03$, $r=0.15$, $p=0.07$, $r=0.09$). Tranwell shows the infiltration of THP-1of **(C)** NCI-H520 cells and **(D)** LTEP-s cells. Western blotting assay showing the expression of stemness marker genes SOX2, Nanog, CD133, CD44, KLF4 following FG G knockdown in **(E)** NCI-H520 and **(F)** LTEP-s cells. "ns" No Significant.

observed in patients with malignant tumors and is closely related to tumor invasion, metastasis (43–45), angiogenesis (46), and tumor growth processes (47); further, its degradation products with carcinogenesis have been reported in tumors (41). However, Nagata et al. found that frameshift mutations in FG G led to hypofibrinemia, indicating that FG G was involved in the regulation of fibrinogen secretion (48). In addition, FG G inhibits platelet adhesion to fibrinogen by interacting with hepatitis B splicing protein (49). Dysregulation of FG G has also been reported in many malignant tumor types, such as liver cancer (50), stomach cancer (40), and prostate cancer (51), underscoring its potential

relevance as a tumor marker. FG G is an important adverse prognostic factor for gastric cancer (35). Another study showed that serum FG G levels predicted the progression of prostate cancer (51). Additionally, FG G is thought to distinguish cancer from normal sera as a potential tumor marker in pancreatic cancer (52). Additional data show the possibility of urine FG G levels as a potential diagnostic marker for NSCLC (53). These findings suggest that FG G could hold diagnostic, prognostic, and therapeutic implications in cancer.

Our bioinformatics modeling demonstrated that FG G as a risk prognosticator is of significant research value in LUSC, and the

results of subsequent *in vitro* experiments are consistent with reports of abnormal expression of FGG mRNA in various cancers. Knockdown of FGG caused functional changes in LUSC tumor progression at the tumor cell level, significantly inhibited the proliferation and clonogenesis ability of NCI-H520 and LTEP-s cells, and blocked the cell cycle in the S phase (Supplementary Figure 9). It also inhibited the invasion and migration ability of tumor cells, by reducing the EMT process and promoting the early apoptosis of tumor cells. In terms of dry characteristics, FGG down-regulation decreased the expressions of KLF4, Nanog, CD44, and SOX2 in NCI-H520 cells, and the expressions of CD44, Nanog, and CD133 in LTEP-s cells. In terms of immune cell infiltration, the expression of FGG in LUSC tissues was significantly correlated with M0 and M1 type macrophages, while knockdown of FGG in LUSC cells significantly affected the degree of immune infiltration of M1 type macrophages (Supplementary Figure 10) formed by polarization of THP-1 cells, suggesting that FGG plays a specific role in the immune infiltration of LUSC.

In summary, our study successfully constructed a LUSC-specific DEGs based risk and prognosis model and verified the reliability of the risk model from the data model. According to the prognostic risk factors, including tumor invasion, metastasis, survival, immune infiltration, and tumor stem cell-related genes, DEGs in LUSC were used to determine associations between functional genes and tumor progression, immune invasion, and dry index. However, this prognostic model has some limitations, such as the relatively simple database and limited factors analyzed (such as lack of smoking status, drivers, ORRs of various chemotherapy treatments, and subsets of immune checkpoint blocking). Subsequently, *in vitro* studies of the LUSC-specific prognostic marker FGG will provide deeper insights into LUSC. As a risk factor in this prognostic model, FGG significantly inhibited the progression of LUSC tumor cells after knockdown and reduced the expression of dry marker genes and the infiltration level of M1 type macrophages, suggesting that FGG is a potential biomarker for independent poor prognosis of LUSC to identify LUSC patients with poor clinical outcomes and that it may play specific roles in dry maintenance and immune infiltration. However, the specific mechanism underlying the changes in tumor progression warrants further study.

5 Conclusion

This study established a seven-gene profile (FGG, C3, FGA, JUN, CST3, CPSF4, and HIST1H2BH) prognostic stratification system demonstrated in LUSC based on Tumor Progression, Immune Infiltration, and Stem Index. *In vitro* experiments confirmed that DEGs FGG could be independent prognostic biomarkers of LUSC promoting cell proliferation, migration, invasion, THP-1 cell infiltration, and stem cell maintenance.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Ethics statement

The current study was approved by the institutional ethics review board of Affiliated Hospital of Inner Mongolia Medical University, Hohhot, Inner Mongolia Autonomous Region, China (NO.KY(2021021)). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

RW contributed to the study's conception, design, and manuscript writing. RM conducted the *in vitro* experiment. XD performed the bioinformatics analysis. JZ and KL generated all the Figures and Tables. LY and MZ contributed to the manuscript editing. RW was responsible for funding acquisition. CW and PL reviewed and approved the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was funded by the Youth Backbone Project of the Affiliated Hospital of Inner Mongolia Medical University (2022NYFYFG010).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fimmu.2023.1236444/full#supplementary-material>

References

- Chen Z, Fillmore CM, Hammerman PS, Kim CF, Wong KK. Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat Rev Cancer* (2014) 14(8):535–46. doi: 10.1038/nrc3775
- Lu J, Wang W, Xu M, Li Y, Chen C, Wang X. A global view of regulatory networks in lung cancer: An approach to understand homogeneity and heterogeneity. *Semin Cancer Biol* (2017) 42:31–8. doi: 10.1016/j.semcancer.2016.11.004
- Reli V, Trerotola M, Guerra E, Alberti S. Abandoning the notion of non-small cell lung cancer. *Trends Mol Med* (2019) 25(7):585–94. doi: 10.1016/j.molmed.2019.04.012
- Ren S, Xiong X, You H, Shen J, Zhou P. The combination of immune checkpoint blockade and angiogenesis inhibitors in the treatment of advanced non-small cell lung cancer. *Front Immunol* (2021) 12:689132. doi: 10.3389/fimmu.2021.689132
- Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA: Cancer J Clin* (2019) 69(1):7–34. doi: 10.3322/caac.21551
- Cancer Genome Atlas Research Network. Comprehensive genomic characterization of squamous cell lung cancers. *Nature* (2012) 489(7417):519–25. doi: 10.1038/nature11404
- Huang G, Huang Q, Xie Z, Zhou H, Cao J, Shi L, et al. A nine-long non-coding RNA signature for prognosis prediction of patients with lung squamous cell carcinoma. *Cancer Biomarkers section A Dis Markers* (2019) 26(3):239–47. doi: 10.3233/CBM-182275
- Duruiseaux M, Esteller M. Lung cancer epigenetics: From knowledge to applications. *Semin Cancer Biol* (2018) 51:116–28. doi: 10.1016/j.semcancer.2017.09.005
- Cao S, Wang Z, Gao X, He W, Cai Y, Chen H, et al. FOXCl induces cancer stem cell-like properties through upregulation of beta-catenin in NSCLC. *J Exp Clin Cancer Res CR* (2018) 37(1):220. doi: 10.1186/s13046-018-0894-0
- Deng K, Liu L, Tan X, Zhang Z, Li J, Ou Y, et al. WIP1 promotes cancer stem cell properties by inhibiting p38 MAPK in NSCLC. *Signal transduction targeted Ther* (2020) 5(1):36. doi: 10.1038/s41392-020-0126-x
- Horvath L, Thienpont B, Zhao L, Wolf D, Pircher A. Overcoming immunotherapy resistance in non-small cell lung cancer (NSCLC) - novel approaches and future outlook. *Mol Cancer* (2020) 19(1):141. doi: 10.1186/s12943-020-01260-z
- Bass AJ, Watanabe H, Mermel CH, Yu S, Perner S, Verhaak RG, et al. SOX2 is an amplified lineage-survival oncogene in lung and esophageal squamous cell carcinomas. *Nat Genet* (2009) 41(11):1238–42. doi: 10.1038/ng.465
- Mao D, Zhou Z, Song S, Li D, He Y, Wei Z, et al. Identification of stemness characteristics associated with the immune microenvironment and prognosis in gastric cancer. *Front Oncol* (2021) 11:626961. doi: 10.3389/fonc.2021.626961
- Liao Y, Xiao H, Cheng M, Fan X. Bioinformatics analysis reveals biomarkers with cancer stem cell characteristics in lung squamous cell carcinoma. *Front Genet* (2020) 11:427. doi: 10.3389/fgene.2020.00427
- Tanaka A, Sakaguchi S. Regulatory T cells in cancer immunotherapy. *Cell Res* (2017) 27(1):109–18. doi: 10.1038/cr.2016.151
- Cassetta L, Pollard JW. Targeting macrophages: therapeutic approaches in cancer. *Nat Rev Drug Discovery* (2018) 17(12):887–904. doi: 10.1038/nrd.2018.169
- Wang Y, Zhang X, Yang L, Xue J, Hu G. Blockade of CCL2 enhances immunotherapeutic effect of anti-PD1 in lung cancer. *J Bone Oncol* (2018) 11:27–32. doi: 10.1016/j.jbo.2018.01.002
- Ahern E, Cubitt A, Ballard E, Teng MWL, Dougall WC, Smyth MJ, et al. Pharmacodynamics of Pre-Operative PD1 checkpoint blockade and receptor activator of NFkB ligand (RANKL) inhibition in non-small cell lung cancer (NSCLC): study protocol for a multicentre, open-label, phase 1B/2, translational trial (POPCORN). *Trials* (2019) 20(1):753. doi: 10.1186/s13063-019-3951-x
- Liu D, Schilling B, Liu D, Sucker A, Livingstone E, Jerby-Arnon L, et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 blockade in patients with metastatic melanoma. *Nat Med* (2019) 25(12):1916–27. doi: 10.1038/s41591-019-0654-5
- Li B, Chan HL, Chen P. Immune checkpoint inhibitors: basics and challenges. *Curr medicinal Chem* (2019) 26(17):3009–25. doi: 10.2174/0929867324666170804143706
- Eun K, Ham SW, Kim H. Cancer stem cell heterogeneity: origin and new perspectives on CSC targeting. *BMB Rep* (2017) 50(3):117–25. doi: 10.5483/BMBRep.2017.50.3.222
- Yang CY, Yang JC, Yang PC. Precision management of advanced non-small cell lung cancer. *Annu Rev Med* (2020) 71:117–36. doi: 10.1146/annurev-med-051718-013524
- Shi YX, Wang Y, Li X, Zhang W, Zhou HH, Yin JY, et al. Genome-wide DNA methylation profiling reveals novel epigenetic signatures in squamous cell lung cancer. *BMC Genomics* (2017) 18(1):901. doi: 10.1186/s12864-017-4223-3
- Chen M, Wang X, Wang W, Gui X, Li Z. Immune- and stemness-related genes revealed by comprehensive analysis and validation for cancer immunity and prognosis and its nomogram in lung adenocarcinoma. *Front Immunol* (2022) 13:829057. doi: 10.3389/fimmu.2022.829057
- Zhang Z, Zhang S, Li X, Zhao Z, Chen C, Zhang J, et al. Reference genome and annotation updates lead to contradictory prognostic predictions in gene expression signatures: a case study of resected stage I lung adenocarcinoma. *Briefings Bioinf* (2021) 22(3):bbaa081. doi: 10.1093/bib/bbaa081
- Li N, Li Y, Zheng P, Zhan X. Cancer stemness-based prognostic immune-related gene signatures in lung adenocarcinoma and lung squamous cell carcinoma. *Front Endocrinol* (2021) 12:755805. doi: 10.3389/fendo.2021.755805
- Liu H, Li T, Dong C, Lyu J. Identification of miRNA signature for predicting the prognostic biomarker of squamous cell lung carcinoma. *PLoS One* (2022) 17(3):e0264645. doi: 10.1371/journal.pone.0264645
- Wang X, Su R, Guo Q, Liu J, Ruan B, Wang G. Competing endogenous RNA (ceRNA) hypothetic model based on comprehensive analysis of long non-coding RNA expression in lung adenocarcinoma. *PeerJ* (2019) 7:e8024. doi: 10.7717/peerj.8024
- Ran R, Jin JW, Zhang WP. MALAT-1 expression correlates with prognosis in non-small-cell lung carcinoma: A systematic review and meta-analysis. *Dis Markers* (2021) 2021:5424623. doi: 10.1155/2021/5424623
- Huang TH, Wu ATH, Cheng TS, Lin KT, Lai CJ, Hsieh HW, et al. In silico identification of thiothrepton as an inhibitor of cancer stem cell growth and an enhancer for chemotherapy in non-small-cell lung cancer. *J Cell Mol Med* (2019) 23(12):8184–95. doi: 10.1111/jcmm.14689
- Herreros-Pomares A, de-Maya-Girones JD, Calabuig-Fariñas S, Lucas R, Martínez A, Pardo-Sánchez JM, et al. Lung tumorspheres reveal cancer stem cell-like properties and a score with prognostic impact in resected non-small-cell lung cancer. *Cell Death Dis* (2019) 10(9):660. doi: 10.1038/s41419-019-1898-1
- Malta TM, Sokolov A, Gentles AJ, Burzykowski T, Poisson L, Weinstein JN, et al. Machine learning identifies stemness features associated with oncogenic dedifferentiation. *Cell* (2018) 173(2):338–354.e315. doi: 10.1016/j.cell.2018.03.034
- Tahmasebi S, Amiri M, Sonenberg N. Translational control in stem cells. *Front Genet* (2018) 9:709. doi: 10.3389/fgene.2018.00709
- Shibue T, Weinberg RA. EMT, CSCs, and drug resistance: the mechanistic link and clinical implications. *Nat Rev Clin Oncol* (2017) 14(10):611–29. doi: 10.1038/nrclinonc.2017.44
- Fukumura D, Kloepper J, Amoozgar Z, Duda DG, Jain RK. Enhancing cancer immunotherapy using antiangiogenics: opportunities and challenges. *Nat Rev Clin Oncol* (2018) 15(5):325–40. doi: 10.1038/nrclinonc.2018.29
- Simurda T, Brunclikova M, Asselta R, Caccia S, Zolkova J, Kolkova Z, et al. Genetic variants in the FGB and FGG genes mapping in the beta and gamma nodules of the fibrinogen molecule in congenital quantitative fibrinogen disorders associated with a thrombotic phenotype. *Int J Mol Sci* (2020) 21(13):4616. doi: 10.3390/ijms21134616
- Yokoyama K, Zhang XP, Medved L, Takada Y. Specific binding of integrin alpha v beta 3 to the fibrinogen gamma and alpha E chain C-terminal domains. *Biochemistry* (1999) 38(18):5872–7. doi: 10.1021/bi9827619
- Yokoyama K, Erickson HP, Ikeda Y, Takada Y. Identification of amino acid sequences in fibrinogen gamma -chain and tenascin C C-terminal domains critical for binding to integrin alpha v beta 3. *J Biol Chem* (2000) 275(22):16891–8. doi: 10.1074/jbc.M000610200
- Akakura N, Hoogland C, Takada YK, Saegusa J, Ye X, Liu FT, et al. The COOH-terminal globular domain of fibrinogen gamma chain suppresses angiogenesis and tumor growth. *Cancer Res* (2006) 66(19):9691–7. doi: 10.1158/0008-5472.CAN-06-1686
- Duan S, Gong B, Wang P, Huang H, Luo L, Liu F. Novel prognostic biomarkers of gastric cancer based on gene expression microarray: COL12A1, GSTA3, FGA and FGG. *Mol Med Rep* (2018) 18(4):3727–36. doi: 10.3892/mmr.2018.9368
- Wang M, Zhang G, Zhang Y, Cui X, Wang S, Gao S, et al. Fibrinogen alpha chain knockout promotes tumor growth and metastasis through integrin-AKT signaling pathway in lung cancer. *Mol Cancer Res MCR* (2020) 18(7):943–54. doi: 10.1158/1541-7786.MCR-19-1033
- Peng HH, Wang JN, Xiao LF, Yan M, Chen SP, Wang L, et al. Elevated serum FGG levels prognosticate and promote the disease progression in prostate cancer. *Front Genet* (2021) 12:651647. doi: 10.3389/fgene.2021.651647
- Lyman GH, Khorana AA. Cancer, clots and consensus: new understanding of an old problem. *J Clin Oncol Off J Am Soc Clin Oncol* (2009) 27(29):4821–6. doi: 10.1200/JCO.2009.22.3032
- Niu JY, Tian T, Zhu HY, Liang JH, Wu W, Cao L, et al. Hyperfibrinogenemia is a poor prognostic factor in diffuse large B cell lymphoma. *Ann Hematol* (2018) 97(10):1841–9. doi: 10.1007/s00277-018-3382-x
- Huang W, Wang S, Zhang H, Zhang B, Wang C. Prognostic significance of combined fibrinogen concentration and neutrophil-to-lymphocyte ratio in patients with resectable non-small cell lung cancer. *Cancer Biol Med* (2018) 15(1):88–96. doi: 10.20892/j.issn.2095-3941.2017.0124
- Shiose S, Hata Y, Noda Y, Sassa Y, Takeda A, Yoshikawa H, et al. Fibrinogen stimulates *in vitro* angiogenesis by choroidal endothelial cells via autocrine VEGF. *Graefe's Arch Exp Ophthalmol = Albrecht von Graefes Archiv fur klinische und experimentelle Ophthalmologie* (2004) 242(9):777–83. doi: 10.1007/s00417-004-0910-2
- Adams GN, Rosenfeldt L, Frederick M, Miller W, Waltz D, Kombrinck K, et al. Colon cancer growth and dissemination relies upon thrombin, stromal PAR-1, and fibrinogen. *Cancer Res* (2015) 75(19):4235–43. doi: 10.1158/0008-5472.CAN-15-0964
- Nagata K, Arai S, Taira C, Sugano M, Honda T, Okumura N. A novel frameshift mutation in the fibrinogen γ C terminal region, FGG c.1169_1170 del AT, leading to

hypofibrinogenemia. *Thromb Res* (2017) 159:82–5. doi: 10.1016/j.thromres.2017.10.002

49. Chen JY, Chen WN, Liu LL, Lin WS, Jiao BY, Wu YL, et al. Hepatitis B spliced protein (HBSP) generated by a spliced hepatitis B virus RNA participates in abnormality of fibrin formation and functions by binding to fibrinogen γ chain. *J Med Virol* (2010) 82(12):2019–26. doi: 10.1002/jmv.21918

50. Gao HJ, Chen YJ, Zuo D, Xiao MM, Li Y, Guo H, et al. Quantitative proteomic analysis for high-throughput screening of differential glycoproteins in hepatocellular carcinoma serum. *Cancer Biol Med* (2015) 12(3):246–54. doi: 10.7497/j.issn.2095-3941.2015.0010.

51. Davalieva K, Kiprijanovska S, Maleva Kostovska I, Stavridis S, Stankov O, Komina S, et al. Comparative proteomics analysis of urine reveals down-regulation

of acute phase response signaling and LXR/RXR activation pathways in prostate cancer. *Proteomes* (2017) 6(1):1. doi: 10.3390/proteomes6010001

52. Bloomston M, Zhou JX, Rosemurgy AS, Frankel W, Muro-Cacho CA, Yeatman TJ. Fibrinogen gamma overexpression in pancreatic cancer identified by large-scale proteomic analysis of serum samples. *Cancer Res* (2006) 66(5):2592–9. doi: 10.1158/0008-5472.CAN-05-3659

53. Zhang W, Gao Z, Zeng G, Xie H, Liu J, Liu N, et al. Clinical significance of urinary plasminogen and fibrinogen gamma chain as novel potential diagnostic markers for non-small-cell lung cancer. *Clinica chimica acta; Int J Clin Chem* (2020) 502:55–65. doi: 10.1016/j.cca.2019.11.022