Prognostic model of endoplasmic reticulum stress-related lncRNAs in lung adenocarcinoma: Construction and validation using WGCNA

Haiyang Li[1,A], Zhenshan Zhao[2,D,F], Jing Li[1,B], Yao Rong[1,C], Aimin Zheng[3,B,C], Menghui Hao[2,C]

[1] Department of Medical Oncology, Kailuan General Hospital, Tangshan, China
[2] Department of Thoracic Surgery, Kailuan General Hospital, Tangshan, China
[3] Oncology Department of Integrated Chinese and Western Medicine, Tangshan People's Hospital, China

A – research concept and design; B – collection and/or assembly of data; C – data analysis and interpretation; D – writing the article; E – critical revision of the article; F – final approval of the article

Advances in Clinical and Experimental Medicine, ISSN 1899–5276 (print), ISSN 2451–2680 (online) *Adv Clin Exp Med*. 2026;35(5):847–861

Address for correspondence
Zhenshan Zhao
E-mail: zhaozhenshan1978@126.com

Funding sources
None declared

Conflict of interest
None declared

Received on March 8, 2025
Reviewed on May 5, 2025
Accepted on August 18, 2025

Published online on May 26, 2026

Abstract

Background. Lung adenocarcinoma (LUAD) ranks among the deadliest malignancies worldwide. The endoplasmic reticulum (ER) stress response plays a critical role in the pathogenesis of various cancers, and long non-coding RNAs (lncRNAs) are known for their regulatory roles in gene expression and disease progression.

Objectives. To construct and validate a prognostic model based on ER stress-related lncRNAs in LUAD.

Materials and methods. The Cancer Genome Atlas (TCGA) and Genotype-Tissue Expression (GTEx) databases were used. Utilizing the Molecular Signatures Database (MSigDB), we identified ER stress-related mRNAs and lncRNAs. Weighted gene co-expression network analysis (WGCNA) was employed to identify genes associated with LUAD prognosis. An lncRNA-based prognostic risk scoring model was constructed using univariate and least absolute shrinkage and selection operator (LASSO) regression analyses and independently validated. Consensus clustering analysis was applied to define risk subgroups, optimizing the risk scoring system. The model's performance was evaluated using receiver operating characteristic (ROC) curves and nomograms. Differentially expressed gene (DEG) and enrichment analyses were performed to investigate the biological relevance of the risk score. Additionally, the relationships between risk scores, immune infiltration, and the tumor microenvironment (TME) were explored.

Results. Using WGCNA, we successfully identified genes strongly associated with ER stress in LUAD prognosis. A prognostic model comprising 13 signature genes was developed, demonstrating robust discrimination between high- and low-risk patients, with the high-risk group exhibiting reduced overall survival (OS). The model's predictive accuracy was confirmed through Kaplan–Meier and ROC analyses. Correlation analysis between risk scores and immune infiltration indicated that the model reflects the immune landscape of LUAD. Subgroup analysis using consensus clustering (C1 and C2) revealed more pronounced differences in immune cell dynamics than the binary risk score classification alone.

Conclusions. This study introduces a novel prognostic model based on the co-expression of ER stress-related lncRNAs in LUAD.

Key words: lung adenocarcinoma, immune infiltration, prognostic model, endoplasmic reticulum stress, long non-coding RNAs

Cite as
Li H, Zhao Z, Li J, Rong Y, Zheng A, Hao M. Prognostic model of endoplasmic reticulum stress-related lncRNAs in lung adenocarcinoma: Construction and validation using WGCNA. *Adv Clin Exp Med*. 2026;35(5):847–861. doi:10.17219/acem/209665

DOI
10.17219/acem/209665

Highlights

- A novel prognostic model based on endoplasmic reticulum (ER) stress-related long non-coding RNAs (lncRNAs) was developed for lung adenocarcinoma (LUAD).
- A 13-lncRNA risk signature effectively stratifies LUAD patients into high- and low-risk groups with distinct survival outcomes.
- DNAJC5B is associated with immune cell infiltration and ER stress regulation, highlighting potential therapeutic relevance in LUAD.
- The prognostic model demonstrated robust predictive performance (AUC > 0.7) across training, testing, and validation cohorts.
- ER stress-related lncRNA signatures may support prognostic assessment and personalized immunotherapy in LUAD.

Background

Lung adenocarcinoma (LUAD), the most common histological subtype of non-small cell lung cancer (NSCLC), is one of the most prevalent malignancies worldwide, accounting for approx. 40% of all lung cancer cases.[1–3] According to the World Health Organization (WHO), lung cancer remains the leading cause of cancer-related mortality globally, with adenocarcinoma representing a substantial proportion of these cases.[3] Lung adenocarcinoma is characterized by marked molecular heterogeneity and biological complexity, which contribute to its poor prognosis, particularly in patients diagnosed at advanced stages. The 5-year survival rate for LUAD remains low, often below 20%, highlighting the urgent need for improved diagnostic and therapeutic strategies. Therefore, a deeper understanding of the molecular mechanisms underlying LUAD and the identification of reliable biomarkers for early detection, targeted intervention, and personalized management are of critical importance.[4]

The endoplasmic reticulum (ER) is a critical intracellular organelle responsible for maintaining protein homeostasis through the regulation of protein synthesis, folding, and transport.[5,6] Endoplasmic reticulum stress occurs when ER protein-folding capacity is disrupted, resulting in the accumulation of misfolded proteins. This stress response activates a series of signaling pathways, including the unfolded protein response (UPR), which aims to restore cellular homeostasis.[7] In the context of LUAD, ER stress has been shown to be closely associated with malignant cellular behaviors, including invasion, metastasis, and resistance to anticancer therapies.[8–10] For example, studies have demonstrated that ER stress can promote epithelial–mesenchymal transition (EMT) in LUAD cells, thereby enhancing their invasive and metastatic potential. Additionally, ER stress has been implicated in the development of therapeutic resistance, representing a major challenge in cancer treatment. Therefore, targeting ER stress has emerged as a promising therapeutic strategy for LUAD.[11]

Long non-coding RNAs (lncRNAs) have attracted increasing attention in recent years due to their critical roles in diverse biological processes, including cellular differentiation, development, and gene expression regulation.[12] In LUAD, dysregulated lncRNA expression has been closely associated with tumor progression, metastasis, and patient prognosis.[13–16] For example, specific lncRNAs have been identified as potential oncogenes or tumor suppressors in LUAD, influencing cell proliferation, apoptosis, and migration. Moreover, lncRNAs can interact with ER stress-related signaling pathways, thereby modulating the biological behavior of tumor cells.[17] This crosstalk between lncRNAs and ER stress pathways suggests that ER stress-related lncRNAs may serve not only as promising biomarkers for LUAD but also as potential therapeutic targets.

Recent advances in molecular biology have revealed that the interplay between ER stress and lncRNAs is more complex than previously recognized. For example, the lncRNA *HOTAIR* has been reported to interact with ER stress-related pathways, promoting tumor progression in LUAD through modulation of key genes involved in the UPR.[18] Similarly, the lncRNA *MALAT1* has been implicated in enhancing the metastatic potential of LUAD cells by stabilizing ER stress-related transcripts.[19] These findings further underscore the potential of lncRNAs as both prognostic biomarkers and therapeutic targets in LUAD.

Furthermore, the tumor microenvironment (TME) plays a critical role in the progression of LUAD and in determining treatment response. Endoplasmic reticulum stress has been shown to modulate the TME by influencing the recruitment, activation, and function of immune cells, including macrophages and T lymphocytes.[20] This interaction between ER stress and the TME further complicates the therapeutic landscape of LUAD and highlights the need for a comprehensive understanding of the underlying molecular mechanisms. The complex interplay among ER stress, lncRNAs, and the TME in LUAD underscores the importance of further investigation into their roles in tumor biology. A deeper understanding of these interactions may reveal novel therapeutic targets and contribute to improved clinical outcomes in LUAD.

Objectives

Although previous studies have explored the roles of ER stress and ER stress-related lncRNAs in LUAD, their interactions and specific prognostic applications in LUAD remain unclear. This study aimed to construct an ER stress-related prognostic risk model through comprehensive analysis of gene expression data from LUAD patients, with particular emphasis on lncRNA expression patterns. We anticipate that this study will provide novel insights and potential biomarkers for personalized prognostic assessment and therapeutic decision-making in LUAD.

Material and methods

Data processing

This study utilized data from The Cancer Genome Atlas (TCGA; https://portal.gdc.cancer.gov) and the Genotype-Tissue Expression (GTEx) project (https://gtexportal.org). In addition, single-cell validation analyses were performed using TISCH2 (http://tisch.comp-genomics.org), which includes 17 single-cell sequencing datasets. TISCH2 is a single-cell transcriptomic database focused on the TME, providing detailed cell-type annotations. As LUAD was not available as a specific category in the TISCH2 database, NSCLC was selected as the closest alternative for the initial validation analysis. This approach enabled validation using the available single-cell transcriptomic data relevant to the lung cancer microenvironment.

The TCGA cohort of patients with LUAD included transcriptomic data and detailed clinical information for 600 patients. All patients underwent surgical resection and had relatively complete follow-up data, with a median follow-up duration of 6 years. The patients ranged in age from 30 to 90 years, with a male-to-female ratio of approx. 1.8:1. The pathological stage distribution included 279 patients with stage I disease, 124 with stage II, 100 with stage IIA, 85 with stage III, and 26 with stage IV disease.

Identification of ER stress-related lncRNAs

A keyword search for "endoplasmic reticulum stress" identified 419 ER stress-related genes. Pearson's correlation analysis was subsequently performed between these ER stress-related genes and 16,876 lncRNAs in the expression matrix, identifying 3,611 lncRNAs significantly associated with ER stress.

Weighted correlation network analysis

Weighted correlation network analysis (WGCNA; v. 1.73) was performed to identify gene modules associated with LUAD prognosis. Selection of an appropriate soft-thresholding power is a critical step in WGCNA, as it directly influences the scale-free topology of the constructed network. To determine the optimal soft threshold, a systematic evaluation of candidate soft-thresholding powers ranging from 1 to 20 was performed. For each candidate value, the scale-free topology fit index (R^2) was calculated to assess the degree of conformity between the network connectivity distribution and the ideal scale-free topology. A soft-thresholding power of 16 was selected, at which the scale-free topology fit index reached 0.80, indicating a satisfactory approximation to a scale-free network model. This threshold was chosen based on the balance between module stability, intramodular connectivity, and reduction of background noise, thereby improving the identification of prognostically relevant gene modules in LUAD.

Construction and validation of the prognostic ER stress-related lncRNA risk score

For construction of the prognostic risk model, differential expression analysis between LUAD and control samples was first performed using the limma package (v. 3.50.3; https://bioconductor.org/packages/limma) in R to identify significantly differentially expressed genes (DEGs). Differentially expressed genes were defined using the thresholds of |log2 fold change| >1 and Benjamini–Hochberg adjusted $p < 0.05$. Subsequently, WGCNA was applied to the gene expression dataset to identify gene modules significantly associated with LUAD prognosis. WGCNA enables the identification of potential disease-related biomarkers by constructing gene co-expression networks and clustering genes with similar expression patterns into biologically relevant modules.

After identifying disease-related gene modules, intersecting analysis was performed between these modules and the previously identified DEGs. This step was used to further refine the candidate gene set by focusing on genes that were both differentially expressed and associated with disease-relevant molecular modules. Subsequently, univariate Cox regression analysis was performed to identify genes significantly associated with patient prognosis among the intersecting genes. Univariate Cox regression is a standard survival analysis method used to evaluate the association between gene expression levels and patient survival outcomes.

To further optimize the prognostic model and reduce the risk of overfitting, least absolute shrinkage and selection operator (LASSO) Cox regression analysis was performed to select prognostically relevant genes and estimate their coefficients. LASSO regression is a regularization-based method suitable for variable selection in high-dimensional datasets. The analysis was conducted using the glmnet package in R (https://CRAN.R-project.org/package=glmnet), and the optimal penalty parameter (λ) was determined through 10-fold cross-validation

by minimizing the partial likelihood deviance. This approach improved model robustness and generalizability.

Validation of prognostic model and consensus clustering

Based on gene expression profiles following risk stratification, consensus clustering analysis was performed on the TCGA-LUAD dataset using the ConsensusClusterPlus package (https://www.bioconductor.org/packages/release/bioc/html/ConsensusClusterPlus.html). Principal component analysis (PCA) and t-distributed stochastic neighbor embedding (t-SNE) were used to visualize the distribution of clustered samples. A heatmap was generated to illustrate the distribution of LUAD patients across different clusters, together with survival information and stratification into high- and low-risk groups. A nomogram was constructed to provide a visual representation of the clinical applicability of the prognostic model. In addition, subgroup analyses based on clinical characteristics were performed to evaluate the prognostic value of the risk score and consensus clustering within specific patient subgroups.

Enrichment and TME analysis

Immune cell infiltration in LUAD samples was initially evaluated using the CIBERSORT algorithm (R script v. 1.03) (https://cibersort.stanford.edu) with 1,000 permutations and a significance threshold of $p < 0.05$, in combination with the IOBR (Immuno-Oncology Biological Research) package (v. 0.99.9; https://CRAN.R-project.org/package=IOBR) in R to estimate the relative abundance of immune cell populations based on RNA transcriptomic data. This approach enabled quantitative assessment of immune cell infiltration across individual samples. For TME analysis, the ESTIMATE algorithm implemented in the IOBR package was used to assess stromal and immune components within the tumor tissue. Default gene signatures were applied for immune and stromal cell estimation, and immune, stromal, and ESTIMATE scores were calculated for each TCGA-LUAD sample. Patients were stratified into high- and low-risk groups, as well as consensus clustering subgroups, using the median risk score as the cutoff. Differences in stromal, immune, and ESTIMATE scores between groups were assessed using the Wilcoxon rank-sum test.

Gene mutation analysis

Somatic mutation data for TCGA-LUAD patients were obtained from the TCGA database. A waterfall plot was generated to visualize the mutation landscape of frequently mutated genes in the high- and low-risk groups. In addition, tumor mutation burden (TMB) was calculated for each TCGA-LUAD patient using the maftools R package (v. 3.21; https://bioconductor.org/packages/release/bioc/html/maftools.html). Tumor mutation burden was defined as the number of somatic mutations per megabase (Mb) of sequenced genomic region and was subsequently used for comparative and correlation analyses.

LncRNA-related target genes

Potential target genes of the model lncRNAs were identified using RNAct (http://rnact.crg.eu) and systematically compiled for downstream analysis. An interaction network was constructed and hub gene analysis was performed using Cytoscape (https://cytoscape.org). Core target genes associated with the model lncRNAs were subsequently subjected to differential expression analysis across risk score groups and consensus clustering subgroups.

Statistical analyses

Statistical analyses and data visualization were performed using R v. 4.1.0 (R Foundation for Statistical Computing, Vienna, Austria). Survival analyses were conducted using the Kaplan–Meier method, and differences between groups were compared using the log-rank test. Associations between risk scores and clinical characteristics were evaluated using Cox proportional hazards regression analysis with the survival package (v. 3.2-11; https://github.com/therneau/survival) in R to assess the independent prognostic value of the risk score. During model construction and validation, the proportional hazards (PH) assumption was assessed for all Cox regression models, including univariable and LASSO Cox regression analyses. The PH assumption was evaluated using the Schoenfeld residuals test, and the results confirmed that the included variables satisfied this assumption (Supplementary Fig. 1).[21] Statistical tests were selected according to data distribution and variance characteristics. For comparisons between 2 groups, Student's t-test was applied when data met assumptions of normality and homogeneity of variance; Welch's t-test was used when normality was satisfied but variances were unequal; otherwise, the Wilcoxon rank-sum test was applied. A $p < 0.05$ was considered statistically significant.

Results

Model construction

Differential expression analysis and WGCNA were performed on LUAD and normal control samples from the TCGA and GTEx datasets (Fig. 1A,B). The intersection between differentially expressed genes and genes from the blue and gray WGCNA modules was selected for univariate Cox regression analysis (Fig. 1D), yielding candidate genes for prognostic model construction. Subsequently, LASSO Cox regression analysis was performed to construct the ER stress-related lncRNA prognostic model (Fig. 1E,F), with the following risk score formula:

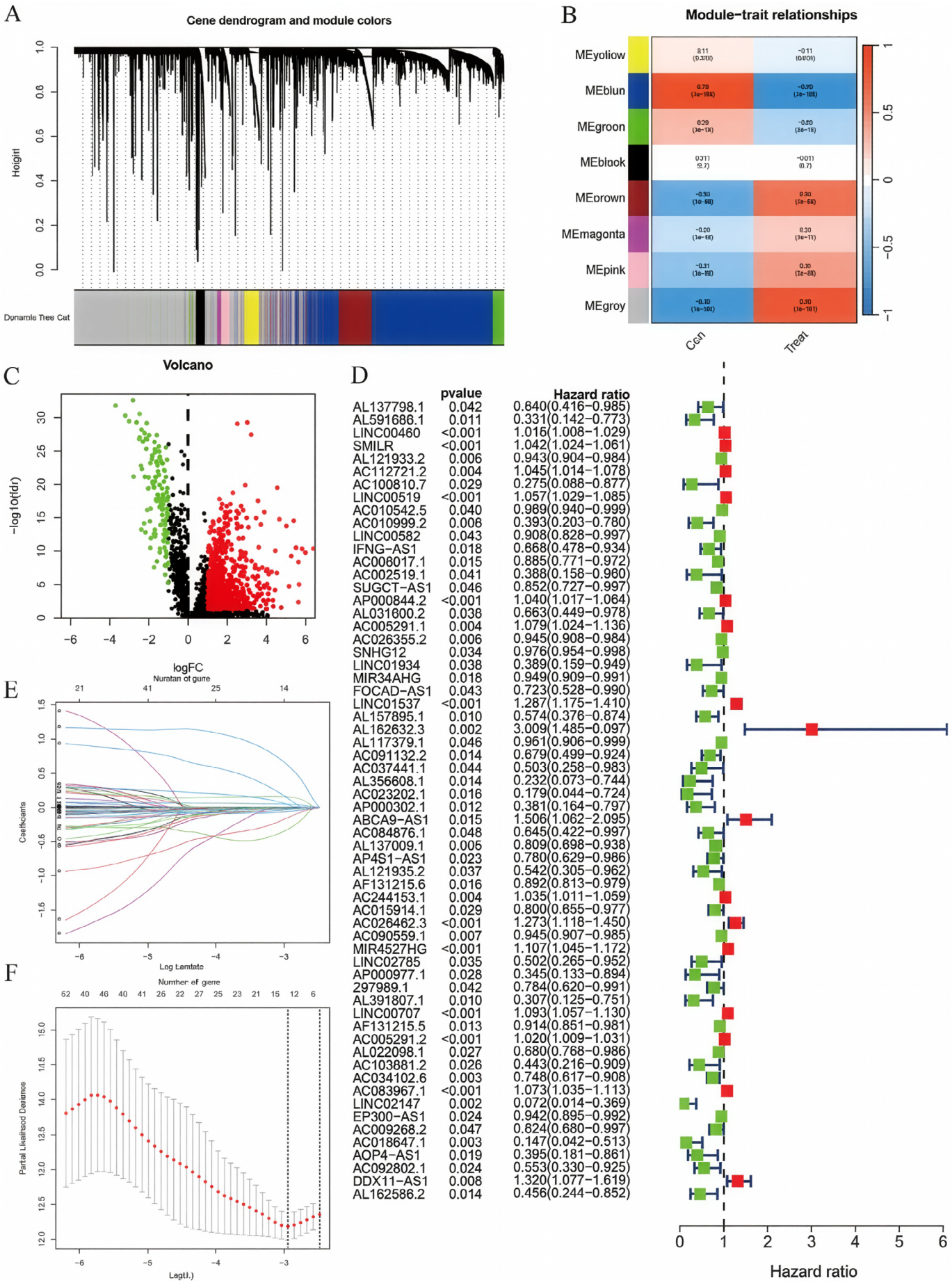

Fig. 1. Endoplasmic reticulum (ER) stress-related long non-coding RNA (lncRNA) model prognostic model. A. Tree-like topology; B. Module-trait correlation heatmap. The horizontal axis represents clinical characteristics, and the vertical axis represents the names of WGCNA modules; C. Volcano plot. Red represents high expression in tumors, and green represents low expression in tumors; D. Univariate Cox forest plot, red represents hazard ratio (HR) >1 and green represents HR < 1; E. Log change coefficients plotted with (least absolute shrinkage and selection operator (LASSO) Cox regression in 10-fold cross-validation; F. Partial likelihood deviation plotted with LASSO Cox regression in 10-fold cross-validation

Risk score =
(*LINC00460* × 0.0006) + (*AC010999.2* × −0.1509) + (*AC026355.2* × −0.0104) + (*LINC01537* × 0.1545) + (*AL162632.3* × 0.6158) + (*ABCA9-AS1* × 0.1352) + (*AL137009.1* × −0.0158) + (*AC026462.3* × 0.0417) + (*AC090559.1* × −0.0023) + (*LINC00707* × 0.0481) + (*AC034102.8* × −0.0575) + (*AC018647.1* × −0.3211) + (*AC092802.1* × −0.0283).
All lncRNAs included in the final prognostic model satisfied the proportional hazards assumption based on Schoenfeld residual testing.

Consensus clustering

Based on the model genes and risk scores, we conducted consensus clustering analysis on LUAD samples. First, the patients were divided into different clusters (K = 2–9), and the optimal number of clusters was determined to be K = 2 based on the consensus matrix (Fig. 2A). The consensus cumulative distribution function (CDF) curve was also plotted (Fig. 2B). The analysis showed that patients with LUAD could be divided into 2 risk groups, C1 and C2. The infiltration level of immunosuppressive cells (such as regulatory T cells, myeloid-derived suppressor cells, and tumor-associated macrophages) in the TME of group C1 (high-risk group) was relatively high. These cells can inhibit the antitumor immune response and promote tumor immune escape. Therefore, group C1 may be associated with a poorer survival prognosis. Conversely, the infiltration level of immunosuppressive cells in group C2 (low-risk group) was lower, suggesting that there may be a more active antitumor immune response in the TME. Effector T cells can recognize and attack tumor cells more effectively, thereby inhibiting tumor growth and metastasis. To visually demonstrate the relationship between the risk score and the clustering results, we plotted a Sankey plot (Fig. 2C) and verified using the Kaplan–Meier curve that the survival time of patients in group C2 was longer (Fig. 2D). Furthermore, the stability and repeatability of the clustering results were confirmed through PCA and t-SNE analysis (Fig. 2E,F).

Model validation and comparison

The samples were randomly divided into a test group and a training group. There were survival differences according to risk scores across the entire cohort, the test set, and the training set, with lower scores indicating better survival outcomes (Fig. 3A–C). Time-dependent receiver operating characteristic (ROC) analysis demonstrated robust predictive accuracy. The area under the ROC curve (AUC) for the risk scores with clinical relevance, as well as the AUCs at 1, 3, and 4 years, were consistent (Fig. 3D,E).

Functional enrichment analysis

Functional enrichment analysis was performed on differential genes between the risk score and consensus clustering groups. Gene Ontology (GO) analysis related to the risk score mainly involved cell movement and ciliary movement. Among these, "cell motility" and "cilia motility" were significantly enriched biological processes that play a key role in LUAD invasion and metastasis (Fig. 4A), while GO analysis related to consensus clustering mainly involved muscle contraction and fiber contraction (Fig. 4B). Kyoto Encyclopedia of Genes and Genomes (KEGG; https://www.genome.jp/kegg) enrichment analysis and gene set enrichment analysis (GSEA) related to the risk score were enriched only in cell movement; therefore, these results are not shown in the figures. The KEGG analysis for consensus clustering mainly involved vascular contraction (Fig. 4C). GSEA for consensus clustering mainly involved the formation and regulation of the stromal group (Fig. 4D).

Immune-related analysis

The risk score and consensus clustering showed similar profiles in immune cell infiltration and immune cell function, primarily associated with T cells and macrophages (Fig. 5A,B,D,E). By integrating checkpoint analysis with risk scores and consensus clustering, together with the analysis of immune cells and their functions, the consensus clustering groups appeared to be more immunologically sensitive (Fig. 5C,F).

Mutation analysis

An in-depth examination of the mutational profile in relation to the risk score was performed. The mutational spectrum revealed the top 20 most frequently mutated genes, highlighting a pronounced disparity in mutation rates between the high- and low-risk cohorts (Fig. 6A). A statistically significant difference in mutation rates was observed between these groups ($p = 0.039$), with the high-risk group exhibiting a notably elevated TMB score compared to the low-risk group (Fig. 6B). Moreover, a significant positive correlation was identified between TMB and risk scores (correlation coefficient $R = 0.087$, $p = 0.033$), suggesting a direct association between the 2 metrics (Fig. 6C).

Clinical and functional exploration of model genes

Pairwise sample analysis was conducted on the model genes, revealing expression differences for all core genes in the model except *AC010999.2*, *AC026462.3*, and *AC034102.8* (Fig. 7A). Prognostic analysis of the model genes over 1–5 years indicated a close relationship between

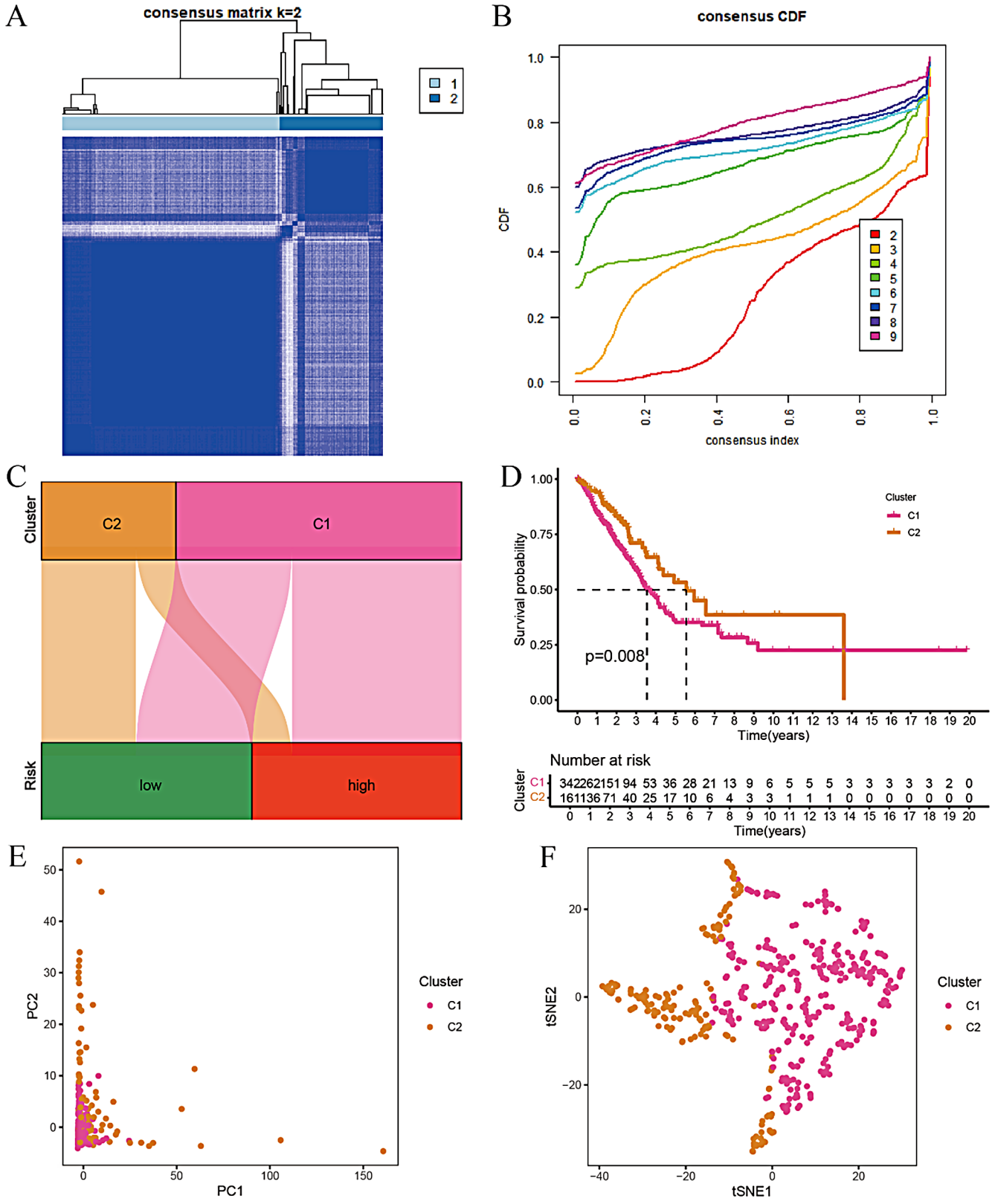

Fig. 2. Identifying risk clusters through consensus clustering. A. Consensus clustering matrix at K = 2; B. Cumulative distribution function (CDF) curves for clustering from k = 2 to 9; C. Sankey diagram; D. Survival analysis of lung adenocarcinoma (LUAD) samples divided into 2 clusters; E. Principal component analysis (PCA) plot for the 2 clusters; F. T-distributed stochastic neighbor embedding (tSNE) plot for the 2 clusters

the model genes and prognosis (Fig. 7B). Immunoinfiltration analysis showed that the model genes were involved in immune cell production and immune cell function (Fig. 7C,D). Based on TCGA data, we further analyzed the expression of the hub target genes in the risk scoring and consensus clustering groups (Fig. 7E,F). Combining

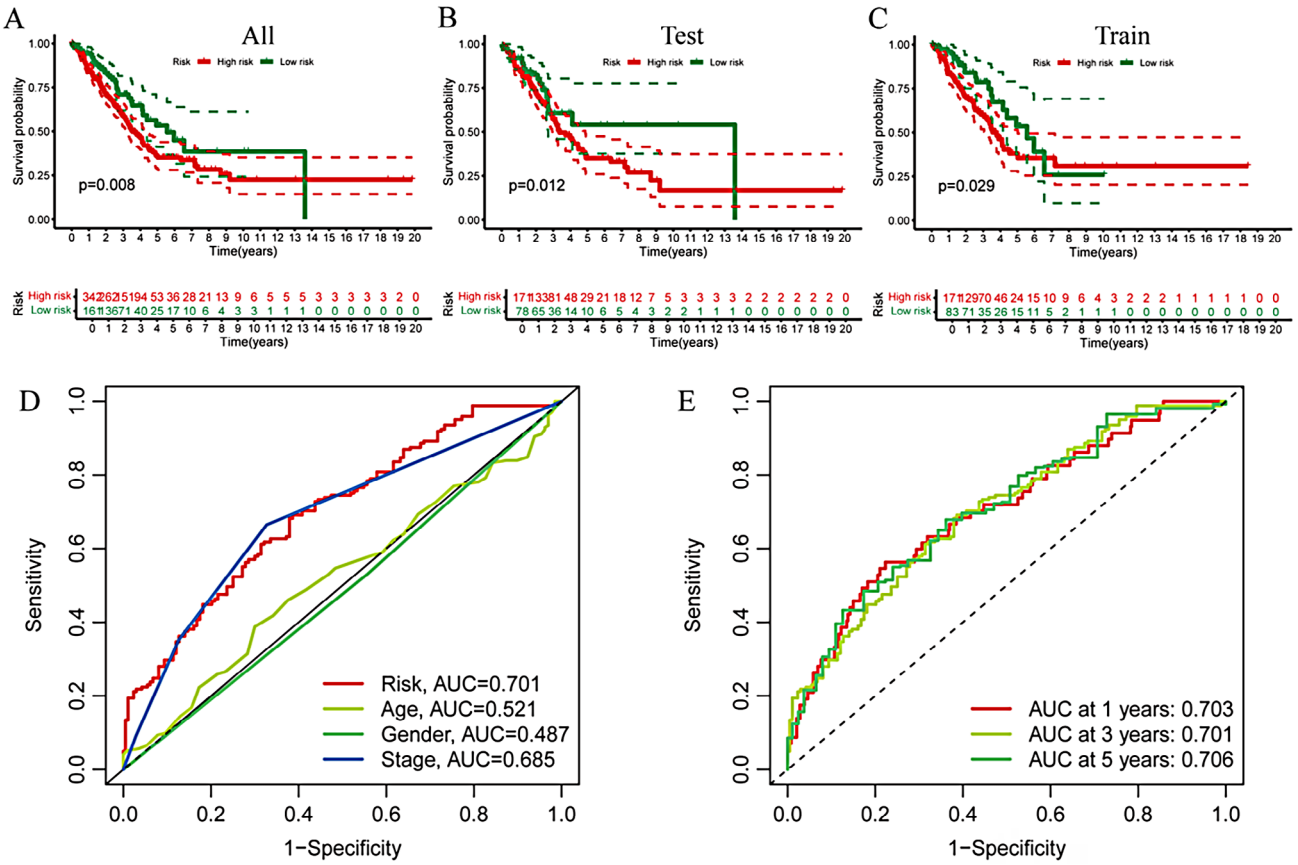

Fig. 3. Validation of the endoplasmic reticulum (ER) stress-related long non-coding RNA (lncRNA) prognostic model. A. Kaplan–Meier survival curves (all); B. Kaplan–Meier survival curves (test); C. Kaplan–Meier survival curves (train); D. Clinically relevant prognostic receiver operating characteristic (ROC); E. Prognostic ROC for risk scores. In the Kaplan–Meier survival curve, red represents the high-risk group and green represents the low-risk group

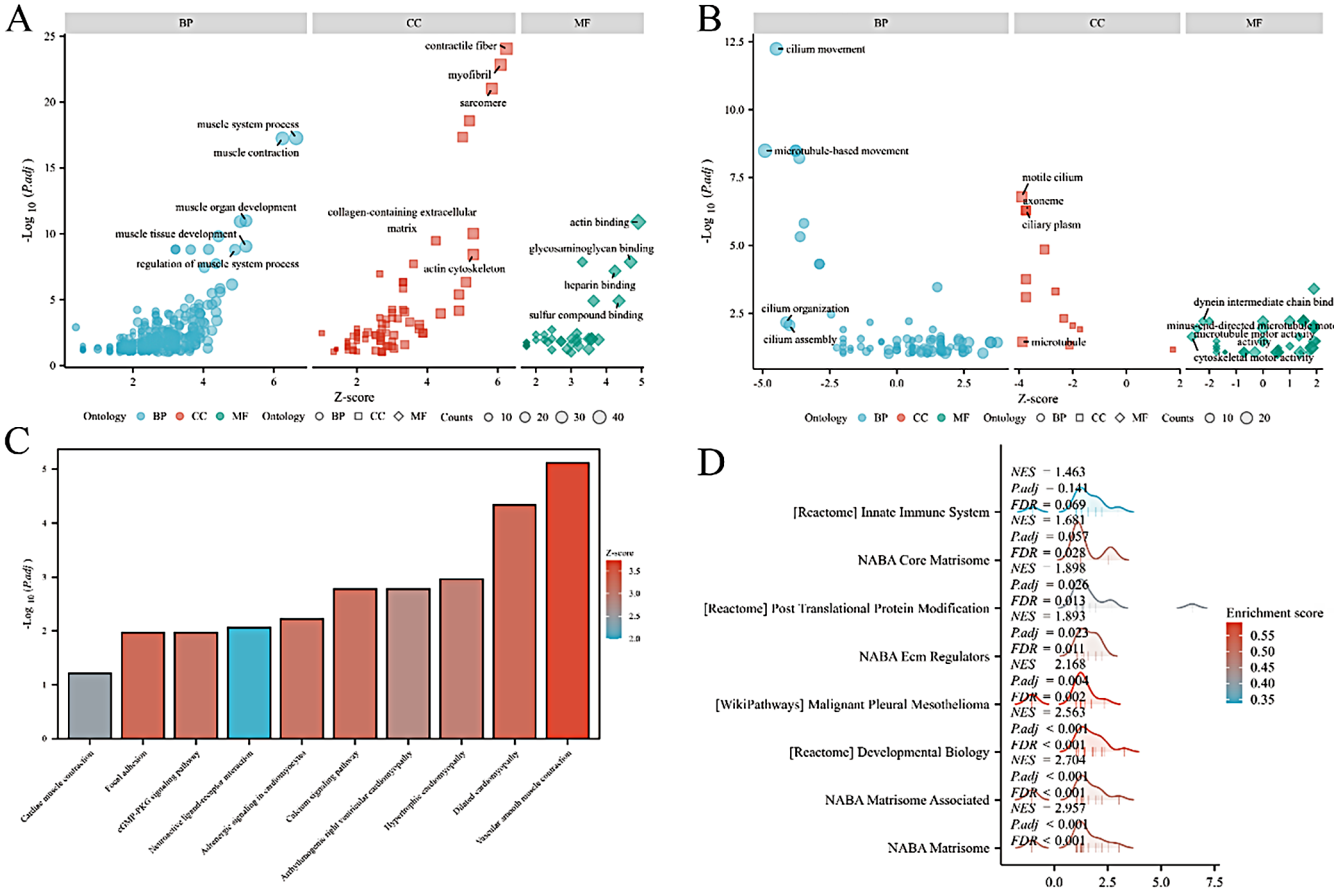

Fig. 4. Functional enrichment analysis for risk score and consensus clustering. A. Gene Ontology (GO) enrichment analysis for the risk score; B. GO enrichment analysis for consensus clustering; C. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis for consensus clustering; D. Gene set enrichment analysis (GSEA) enrichment analysis for consensus clustering

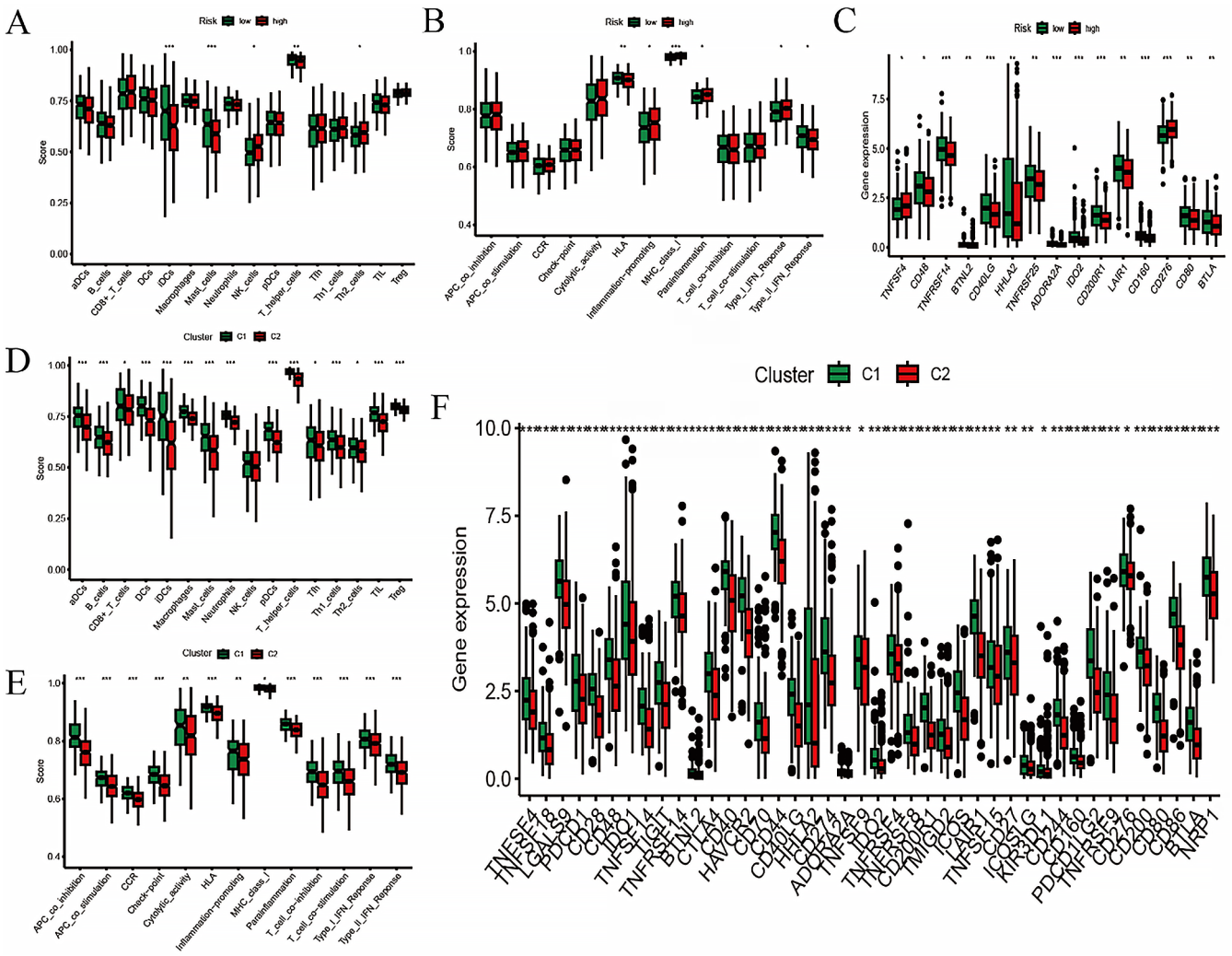

Fig. 5. Immune-related analysis for risk score and consensus clustering. A. Differential analysis of immune cells in risk scoring; B. Differential analysis of immune cell functions in risk scoring; C. Differential analysis of checkpoints in risk scoring; D. Differential analysis of immune cells in consensus clustering; E. Differential analysis of immune cell functions in consensus clustering; F. Differential analysis of checkpoints in consensus clustering. The p-values are represented as *p < 0.05, **p < 0.01, and ***p < 0.001. Green represents C1 or the low-risk group, and red represents C2 or the high-risk group

expression correlation analysis and differential expression analysis, *DNAJC5B* was ultimately selected for subsequent validation. Based on TISCH2, as there was no option for LUAD, we selected NSCLC. *DNAJC5B* was primarily associated with monocytes/macrophages, followed by B cells and natural killer (NK) cells (Fig. 8).

Discussion

The individual variability among patients leads to different responses to treatment, necessitating robust prognostic markers for personalized prediction and therapy.[4,22] Therefore, the construction of a robust prognostic model to predict patient outcomes is urgently needed. Long non-coding RNAs play a significant role in many biological processes, including tumorigenesis, cell differentiation, and metabolism.[23–25] In recent years, these novel non-coding transcripts have garnered widespread attention due to their extensive and complex roles in cancer migration and progression.[26]

Long non-coding RNAs play a key role in cancer development by interacting with ER stress-related proteins, activating downstream signaling pathways, and regulating apoptosis and survival. For example, lncRNA *GAS5* binds to *GRP78*, activates the UPR and the caspase-9 and CHOP signaling pathways, and induces apoptosis in hepatoblastoma HepG2 cells.[27] lncRNA *MEG3* is associated with increased expression of ER stress pathway proteins (e.g., GRP78, PERK, IRE1α, and ATF6), enhances NF-κB (nuclear factor kappa-light-chain-enhancer of activated B cells) translocation, and induces apoptosis in cancer cells.[28–30] In addition, lncRNAs are involved in the fine regulation of apoptosis and survival mechanisms and influence the sensitivity of tumor cells to chemotherapeutic agents.[31,32] These mechanisms are complex and diverse, providing potential targets for cancer prognostic assessment and treatment, and further research is needed to support precision cancer therapy in the future.

In the present study, we paid special attention to the role of lncRNAs such as *LINC00460* in ER stress.

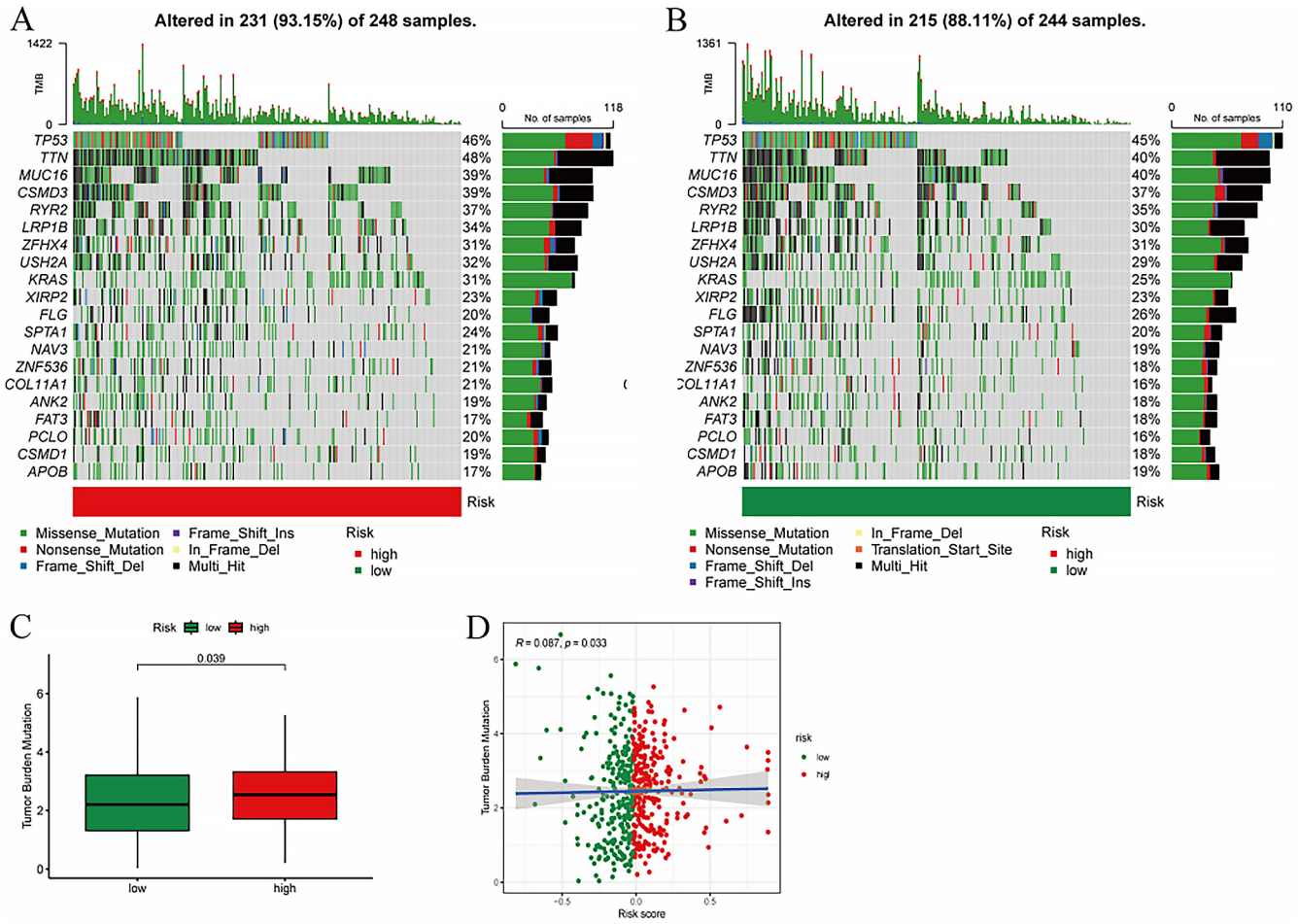

Fig. 6. Exploration of tumor mutation burden based on risk score. A. Mutation landscape of the top 20 genes related to the risk score (high risk); B. Mutation landscape of the top 20 genes related to the risk score (low risk); C. Box plot of tumor mutation burden (TMB) scores for the 2 risk groups; D. Correlation analysis between TMB score and risk score

LINC00460 has been reported to interact with key regulators of ER stress, such as ATF4, which in turn affects UPR signaling. Our data indicate that the expression level of *LINC00460* is closely associated with the prognosis of LUAD patients, suggesting that it may influence tumor progression by regulating the UPR.

We employed the WGCNA method to identify gene modules highly related to the disease from the gene expression data of LUAD patients. Combined with DEGs and univariate Cox analysis, we selected genes significantly associated with prognosis and further optimized model construction using the LASSO Cox regression algorithm. This process not only improved the predictive accuracy of the model but also ensured its robustness through cross-validation.

The risk scoring model and its consensus clustering subgroups demonstrated excellent discrimination ability in both the training and testing sets, clearly distinguishing high- and low-risk patient groups. The predictive accuracy of the model was further validated through Kaplan–Meier survival curves and ROC curve analysis, providing a potential prognostic tool for personalized treatment of LUAD patients. Compared with previous methods that only used the median or average value of the risk score as the threshold to distinguish high and low-risk groups, consensus clustering subgroups (C1 and C2) more accurately combined the risk score with clinical actual subgroups. This combination helps to identify subgroups sensitive to the risk score (C2) and subgroups with unique immune characteristics (C1), providing a more precise stratification basis for clinical treatment. We found that a high risk score was associated with a poorer survival prognosis, which may be related to the high infiltration of immunosuppressive cells in the TME of patients in the high risk score group. This finding provides a potential biomarker for future immunotherapy, suggesting that we can improve patient prognosis by modulating immune cells in the TME.

Furthermore, functional enrichment analysis using the GO and KEGG databases revealed that differential genes were mainly enriched in biological processes related to cell movement and ciliary movement. In contrast, consensus clustering analysis revealed gene enrichment related to muscle contraction and fiber contraction. These findings suggest that ER stress-related lncRNAs may affect the development of LUAD by regulating these key biological processes.

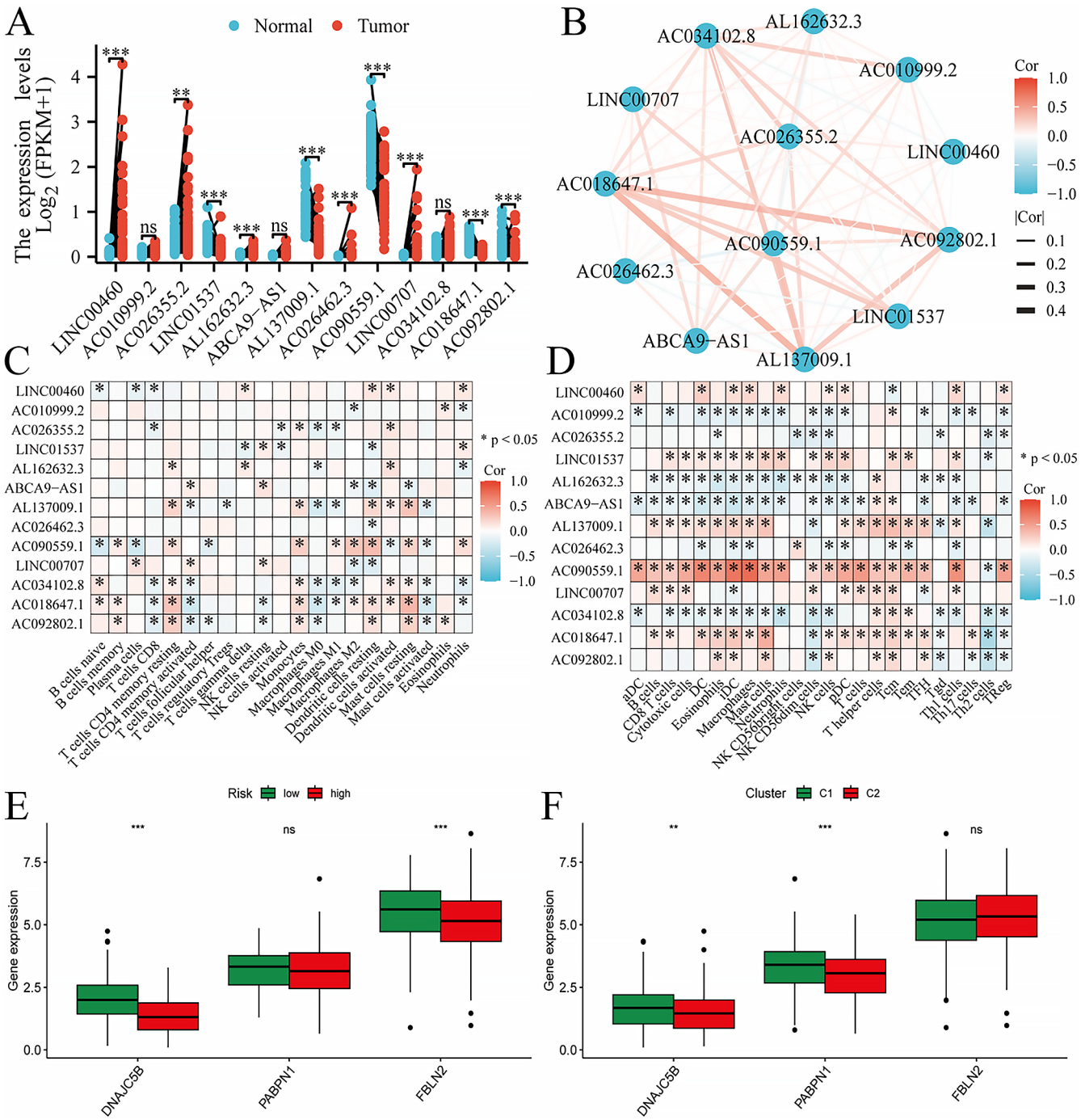

Fig. 7. Clinical and functional exploration of model genes. A. Pairwise differential analysis; B. Correlation network of model genes and target genes; C. Cell and gene correlation (CIBERSORT algorithm); D. Cell and gene correlation (SSGSEA algorithm); E. Differential analysis of core target genes between risk score groups; F. Differential analysis of core target genes between consensus clustering groups. The p-values are represented as *p < 0.05, **p < 0.01, and ***p < 0.001

Muscle contraction and fiber contraction play important roles in tumor biology, as they can regulate vascular contraction and, through the regulation of material exchange between the tumor and the circulatory system, affect tumor metabolism and proliferation on the one hand.[33–35] On the other hand, these processes also participate in the regulation of tumor invasion, metastasis, and immune escape mechanisms.[34,36–38] At the same time, cell movement and ciliary movement directly participate in affecting the invasiveness, metastatic ability, and immune escape mechanisms of tumors.[39,40] It is worth noting that both muscle and fiber contraction and cell and ciliary movement depend on the normal function of the ER to ensure the correct synthesis and execution of protein functions.[41–43] The ER plays a crucial role in maintaining cell movement and tumor biological behavior.

Therefore, the analysis results of this study highlight the multifaceted role of the ER in tumor biology and

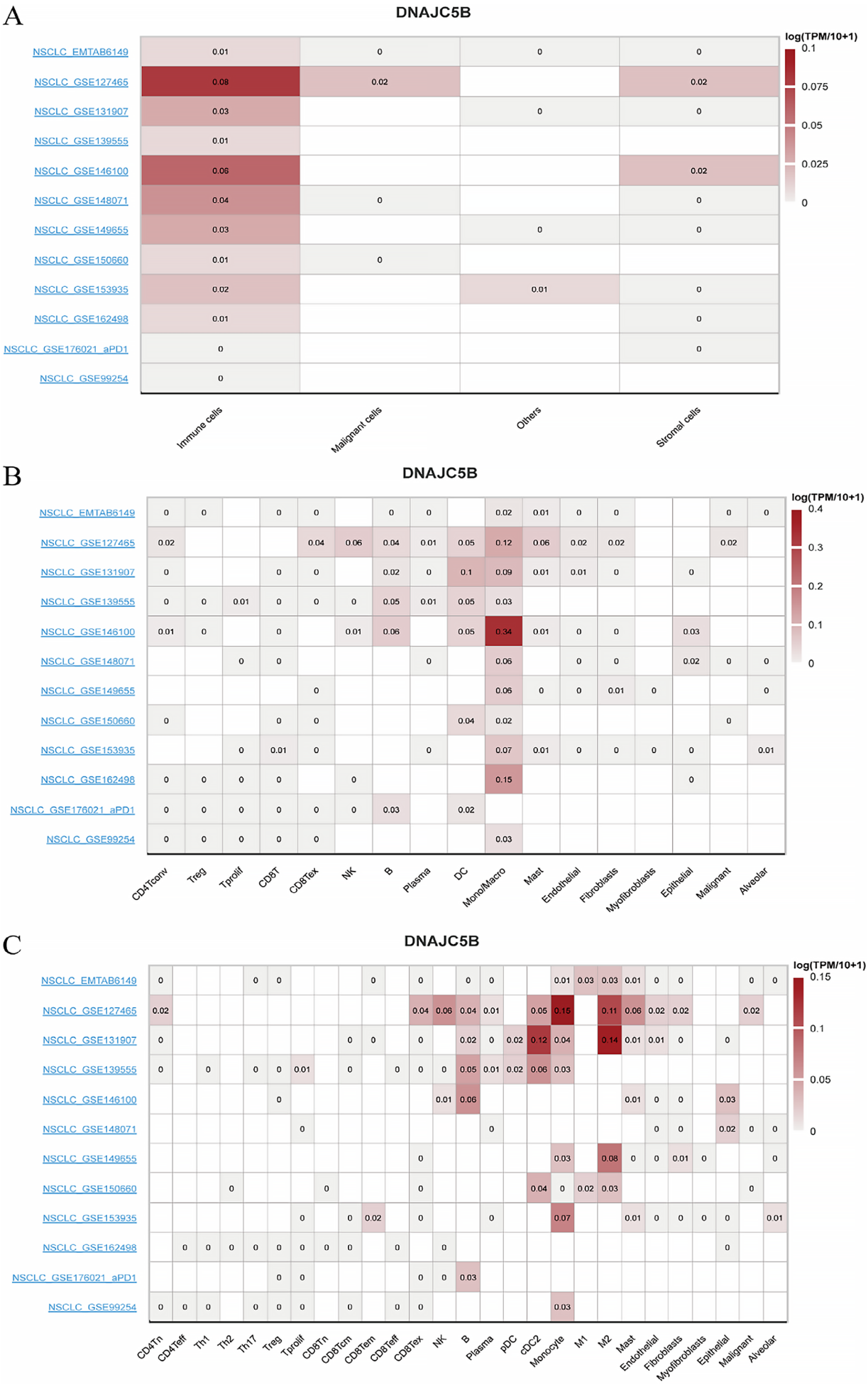

Fig. 8. Single-cell validation. A. Heatmap of malignant tumor groups; B. Heatmap of primary tumor lineage groups; C. Heatmap of secondary tumor lineage groups

provide new perspectives for future research. Especially when developing treatment strategies for LUAD, considering the potential impact of ER stress and related lncRNAs is crucial. These findings may help design new intervention measures to regulate ER stress, thereby affecting tumor development and response to treatment.

The TME, immune system interactions, and TMB are pivotal areas of interest in oncology, significantly influencing tumor genesis and progression.[44–46] Our research demonstrated that the TME score for patients categorized in the high-risk group markedly surpassed that of the low-risk group, suggesting a potential enhanced responsiveness to immunotherapy among high-risk patients. This notion is further corroborated by a detailed examination of immune cell dynamics and their respective roles. Clinically, an elevated TMB score is often linked to enduring clinical benefits and enhanced objective response rates in cancer immunotherapy.[47] Our findings indicate that the TMB score for the high-risk group was considerably higher than for the low-risk group, hinting at a more favorable prognosis and a greater potential for therapeutic advantage from immunotherapy for those with elevated TMB scores.[48]

A combined analysis of risk score grouping and consensus clustering subgroups was performed.[49,50] This finding provides a new perspective for understanding the role of ER stress in the tumor immune microenvironment and may help develop precision treatment strategies for specific immune cell populations.

DNAJC5B (DnaJ heat shock protein family C member 5B), as a member of the heat shock protein family, belongs to the subfamily of DNAJ/Hsp40 protein co-chaperones. These proteins play important roles in cells, including promoting the correct folding of proteins, preventing the aggregation of misfolded proteins, and participating in the degradation and translocation processes of proteins.[51] *DNAJC5B*, with its molecular chaperone characteristics, may be involved in the correct folding of proteins in the ER, preventing the accumulation of abnormal proteins, thereby maintaining protein homeostasis.

When protein folding encounters obstacles, it may trigger ER stress. In this process, *DNAJC5B* may participate in the regulation of the ER stress response through its role in the ER, affecting the cell's response to ER stress. Therefore, *DNAJC5B* may indirectly affect the proliferation, migration, invasion, apoptosis, and colony formation of tumor cells by participating in the ER stress response.[52–54] To verify this hypothesis, we conducted an in-depth exploration of the function of *DNAJC5B* in the experimental part.

In addition, combined with the single-cell sequencing exploration in this study, we found that *DNAJC5B* mainly affects monocytes/macrophages, T cells (NK cells), and B cells. Therefore, we speculate that *DNAJC5B* affects immune cells (monocytes/macrophages, T cells (NK cells), B cells) through the ER stress response, thereby affecting the immune-related treatment of LUAD. Interestingly, the correlation between *DNAJC5B* and immune cells is not related to lung cancer treated with neoadjuvant PD-1 blockade, which provides evidence for our hypothesis.

This study successfully constructed a risk scoring model based on ER stress-related lncRNAs, providing a new perspective and tool for the prognostic evaluation and personalized treatment of LUAD. Through consensus clustering analysis, not only were populations potentially sensitive to treatment identified, but populations with specific immune characteristics were also discovered. In addition, the biological relevance of the core target gene *DNAJC5B* in tumor cell biological behavior was experimentally verified, further confirming the biological relevance of the model.

Limitations of the study

A critical limitation is the absence of external validation cohorts, which may restrict the generalizability of the model. Future studies should prioritize multicenter prospective trials to assess its clinical utility across diverse populations. Additionally, the model's dependency on RNA-seq data necessitates standardization of sequencing platforms for real-world application. The present study's used only data on LUAD patients from the TCGA database, which may be subject to selection bias. In addition, our model has not been validated in an independent cohort, so its generalizability and reliability need to be further assessed. To overcome these limitations, future research will verify the biological function of the model through experimental methods and plans to further explore the clinical application potential of the model in a larger independent patient cohort. Meanwhile, future studies should include validation of our risk scoring model in a larger multicenter cohort of patients and explore its application in guiding individualized treatment. This will help improve the generalizability and reliability of the models, bridge the gap between computational prediction and clinical implementation, and provide more accurate guidance for the clinical treatment of LUAD patients. In addition, we suggest further studies on the regulatory mechanisms of lncRNAs in ER stress and how they affect the immune microenvironment in LUAD.

Conclusions

Our study contributes to the development of new therapeutic strategies to improve outcomes for LUAD patients. This will help improve the model's universality and reliability, providing more accurate guidance for the clinical treatment of LUAD patients. Through these efforts, we hope to translate research results into practical clinical applications, benefiting more LUAD patients.

Supplementary data

The supplementary materials are available at https://doi.org/10.5281/zenodo.16993863. The package contains the following files:

Supplementary File 1. Schoenfeld PH test results.

Data Availability Statement

The datasets supporting the conclusions of this study are publicly available on Figshare at the following link: https://figshare.com/s/0423b4e4fcde57b9dc87.

Consent for publication

All authors have reviewed the manuscript and consent to its publication.

Use of AI and AI-assisted technologies

ORCID iDs

Haiyang Li https://orcid.org/0009-0007-2920-7864
Zhenshan Zhao https://orcid.org/0009-0000-4619-841
Jing Li https://orcid.org/0009-0006-2292-4943
Yao Rong https://orcid.org/0009-0007-1372-1877
Aimin Zheng https://orcid.org/0009-0000-7404-9067
Menghui Hao https://orcid.org/0009-0007-3790-1741

References

1. Cao M, Li H, Sun D, Chen W. Cancer burden of major cancers in China: A need for sustainable actions. *Cancer Commun (Lond)*. 2020;40(5):205–210. doi:10.1002/cac2.12025
2. Ferlay J, Colombet M, Soerjomataram I, et al. Cancer incidence and mortality patterns in Europe: Estimates for 40 countries and 25 major cancers in 2018. *Eur J Cancer*. 2018;103:356–387. doi:10.1016/j.ejca.2018.07.005
3. Travis WD, Brambilla E, Nicholson AG, et al. The 2015 World Health Organization Classification of Lung Tumors. *J Thorac Oncol*. 2015;10(9):1243–1260. doi:10.1097/JTO.0000000000000630
4. Nicholson AG, Tsao MS, Beasley MB, et al. The 2021 WHO Classification of Lung Tumors: Impact of advances since 2015. *J Thorac Oncol*. 2022;17(3):362–387. doi:10.1016/j.jtho.2021.11.003
5. Schwarz DS, Blower MD. The endoplasmic reticulum: Structure, function and response to cellular signaling. *Cell Mol Life Sci*. 2016;73(1):79–94. doi:10.1007/s00018-015-2052-6
6. Balch WE, Morimoto RI, Dillin A, Kelly JW. Adapting proteostasis for disease intervention. *Science*. 2008;319(5865):916–919. doi:10.1126/science.1141448
7. Chen X, Shi C, He M, Xiong S, Xia X. Endoplasmic reticulum stress: Molecular mechanism and therapeutic targets. *Sig Transduct Target Ther*. 2023;8(1):352. doi:10.1038/s41392-023-01570-w
8. Fang C, Weng T, Hu S, et al. IFN-γ-induced ER stress impairs autophagy and triggers apoptosis in lung cancer cells. *OncoImmunology*. 2021;10(1):1962591. doi:10.1080/2162402X.2021.1962591
9. Wan L, Chen Z, Yang J, et al. Identification of endoplasmic reticulum stress-related signature characterizes the tumor microenvironment and predicts prognosis in lung adenocarcinoma. *Sci Rep*. 2023;13(1):19462. doi:10.1038/s41598-023-45690-3
10. Zhu C, Xie Y, Li Q, et al. CPSF6-mediated XBP1 3′UTR shortening attenuates cisplatin-induced ER stress and elevates chemo-resistance in lung adenocarcinoma. *Drug Resist Updat*. 2023;68:100933. doi:10.1016/j.drup.2023.100933
11. Shergalis AG, Hu S, Bankhead A, Neamati N. Role of the ERO1-PDI interaction in oxidative protein folding and disease. *Pharmacol Ther*. 2020;210:107525. doi:10.1016/j.pharmthera.2020.107525
12. Statello L, Guo CJ, Chen LL, Huarte M. Gene regulation by long non-coding RNAs and its biological functions. *Nat Rev Mol Cell Biol*. 2021;22(2):96–118. doi:10.1038/s41580-020-00315-9
13. Lai X, Zhong J, Zhang A, Zhang B, Zhu T, Liao R. Focus on long non-coding RNA MALAT1: Insights into acute and chronic lung diseases. *Front Genet*. 2022;13:1003964. doi:10.3389/fgene.2022.1003964
14. Mondal P, Natesh J, Kamal MA, Meeran SM. Non-coding RNAs in lung cancer chemoresistance. *Curr Drug Metab*. 2020;20(13):1023–1032. doi:10.2174/1389200221666200106105201
15. Inamura K. Major tumor suppressor and oncogenic non-coding RNAs: Clinical relevance in lung cancer. *Cells*. 2017;6(2):12. doi:10.3390/cells6020012
16. Wei H, Zhang S, Lin X, Fang R, Li L. Differential expression and clinical significance of long non-coding RNAs in the development and progression of lung adenocarcinoma. *Front Oncol*. 2024;14:1411672. doi:10.3389/fonc.2024.1411672
17. Ebrahimi N, Saremi J, Ghanaatian M, et al. The role of endoplasmic reticulum stress in the regulation of long noncoding RNAs in cancer. *J Cell Physiol*. 2022;237(10):3752–3767. doi:10.1002/jcp.30846
18. Li L, Liu B, Wapinski OL, et al. Targeted disruption of hotair leads to homeotic transformation and gene derepression. *Cell Rep*. 2013;5(1):3–12. doi:10.1016/j.celrep.2013.09.003
19. Tripathi V, Ellis JD, Shen Z, et al. The nuclear-retained noncoding RNA MALAT1 regulates alternative splicing by modulating SR splicing factor phosphorylation. *Mol Cell*. 2010;39(6):925–938. doi:10.1016/j.molcel.2010.08.011
20. Gonzalez H, Hagerling C, Werb Z. Roles of the immune system in cancer: From tumor initiation to metastatic progression. *Genes Dev*. 2018;32(19–20):1267–1284. doi:10.1101/gad.314617.118
21. Zeng Z, Gao Y, Li J, et al. Violations of proportional hazard assumption in Cox regression model of transcriptomic data in TCGA pan-cancer cohorts. *Comput Struct Biotechnol J*. 2022;20:496–507. doi:10.1016/j.csbj.2022.01.004
22. Ruiz-Cordero R, Devine WP. Targeted therapy and checkpoint immunotherapy in lung cancer. *Surg Pathol Clin*. 2020;13(1):17–33. doi:10.1016/j.path.2019.11.002
23. Tan Y, Lin J, Li T, Li J, Xu R, Ju H. LncRNA-mediated posttranslational modifications and reprogramming of energy metabolism in cancer. *Cancer Commun (Lond)*. 2021;41(2):109–120. doi:10.1002/cac2.12108
24. McCabe EM, Rasmussen TP. lncRNA involvement in cancer stem cell function and epithelial–mesenchymal transitions. *Semin Cancer Biol*. 2021;75:38–48. doi:10.1016/j.semcancer.2020.12.012
25. Yan H, Bu P. Non-coding RNA in cancer. *Essays Biochem*. 2021;65(4):625–639. doi:10.1042/EBC20200032
26. Winkle M, El-Daly SM, Fabbri M, Calin GA. Noncoding RNA therapeutics: Challenges and potential solutions. *Nat Rev Drug Discov*. 2021;20(8):629–651. doi:10.1038/s41573-021-00219-z
27. Zhang W, Zhan H, Li M, Wu G, Liu Z, Wu L. Long noncoding RNA Gas5 induces cell apoptosis and inhibits tumor growth via activating the CHOP-dependent endoplasmic reticulum stress pathway in human hepatoblastoma HepG2 cells. *J Cell Biochem*. 2022;123(2):231–247. doi:10.1002/jcb.30159
28. Zhang Y, Wu J, Jing H, Huang G, Sun Z, Xu S. Long noncoding RNA *MEG3* inhibits breast cancer growth via upregulating endoplasmic reticulum stress and activating NF-κB and p53. *J Cell Biochem*. 2019;120(4):6789–6797. doi:10.1002/jcb.27982
29. Huang ZL, Chen RP, Zhou XT, et al. Long non-coding RNA MEG3 induces cell apoptosis in esophageal cancer through endoplasmic reticulum stress. *Oncol Rep*. 2017;37(5):3093–3099. doi:10.3892/or.2017.5568
30. Chen RP, Huang ZL, Liu LX, et al. Involvement of endoplasmic reticulum stress and p53 in lncRNA MEG3-induced human hepatoma HepG2 cell apoptosis. *Oncol Rep*. 2016;36(3):1649–1657. doi:10.3892/or.2016.4919
31. Ding Z, Kang J, Yang Y. Long non-coding RNA CASC2 enhances irradiation-induced endoplasmic reticulum stress in NSCLC cells through PERK signaling. *3 Biotech*. 2020;10(10):449. doi:10.1007/s13205-020-02443-7
32. Tsuchiya H, Shinonaga R, Sakaguchi H, Kitagawa Y, Yoshida K. NEAT1–SOD2 axis confers sorafenib and lenvatinib resistance by activating AKT in liver cancer cell lines. *Curr Issues Mol Biol*. 2023;45(2):1073–1085. doi:10.3390/cimb45020071

33. Kuczynski EA, Vermeulen PB, Pezzella F, Kerbel RS, Reynolds AR. Vessel co-option in cancer. *Nat Rev Clin Oncol*. 2019;16(8):469–493. doi:10.1038/s41571-019-0181-9
34. Rajabi S, Dehghan MH, Dastmalchi R, Jalali Mashayekhi F, Salami S, Hedayati M. The roles and role-players in thyroid cancer angiogenesis. *Endocr J*. 2019;66(4):277–293. doi:10.1507/endocrj.EJ18-0537
35. Pieterse Z, Sinha D, Kaur P. Pericytes in metastasis. In: Birbrair A, ed. *Pericyte Biology in Disease*. Vol. 1147. Advances in Experimental Medicine and Biology. Cham, Switzerland: Springer International Publishing; 2019:125–135. doi:10.1007/978-3-030-16908-4_5
36. Kuczynski EA, Reynolds AR. Vessel co-option and resistance to anti-angiogenic therapy. *Angiogenesis*. 2020;23(1):55–74. doi:10.1007/s10456-019-09698-6
37. Lugassy C, Vermeulen PB, Ribatti D, Pezzella F, Barnhill RL. Vessel co-option and angiotropic extravascular migratory metastasis: A continuum of tumour growth and spread? *Br J Cancer*. 2022;126(7):973–980. doi:10.1038/s41416-021-01686-2
38. Li S, Zhang Q, Hong Y. Tumor vessel normalization: A window to enhancing cancer immunotherapy. *Technol Cancer Res Treat*. 2020;19:1533033820980116. doi:10.1177/1533033820980116
39. Zanotelli MR, Zhang J, Reinhart-King CA. Mechanoresponsive metabolism in cancer cell migration and metastasis. *Cell Metab*. 2021;33(7):1307–1321. doi:10.1016/j.cmet.2021.04.002
40. Marcadis AR, Kao E, Wang Q, et al. Rapid cancer cell perineural invasion utilizes amoeboid migration. *Proc Natl Acad Sci U S A*. 2023;120(17):e2210735120. doi:10.1073/pnas.2210735120
41. Weigel AV, Chang CL, Shtengel G, et al. ER-to-Golgi protein delivery through an interwoven, tubular network extending from ER. *Cell*. 2021;184(9):2412–2429.e16. doi:10.1016/j.cell.2021.03.035
42. Battson ML, Lee DM, Gentile CL. Endoplasmic reticulum stress and the development of endothelial dysfunction. *Am J Physiol Heart Circ Physiol*. 2017;312(3):H355–H367. doi:10.1152/ajpheart.00437.2016
43. Nieblas B, Pérez-Treviño P, García N. Role of mitochondria-associated endoplasmic reticulum membranes in insulin sensitivity, energy metabolism, and contraction of skeletal muscle. *Front Mol Biosci*. 2022;9:959844. doi:10.3389/fmolb.2022.959844
44. Xiao Y, Yu D. Tumor microenvironment as a therapeutic target in cancer. *Pharmacol Ther*. 2021;221:107753. doi:10.1016/j.pharmthera.2020.107753
45. Kiely M, Lord B, Ambs S. Immune response and inflammation in cancer health disparities. *Trends Cancer*. 2022;8(4):316–327. doi:10.1016/j.trecan.2021.11.010
46. Samstein RM, Lee CH, Shoushtari AN, et al. Tumor mutational load predicts survival after immunotherapy across multiple cancer types. *Nat Genet*. 2019;51(2):202–206. doi:10.1038/s41588-018-0312-8
47. Allgäuer M, Budczies J, Christopoulos P, et al. Implementing tumor mutational burden (TMB) analysis in routine diagnostics: A primer for molecular pathologists and clinicians. *Transl Lung Cancer Res*. 2018;7(5):703–715. doi:10.21037/tlcr.2018.08.14
48. Zhang Y, Wang L, Li R, Liu B. The emerging development of tumor mutational burden in patients with NSCLC. *Future Oncol*. 2020;16(9):469–481. doi:10.2217/fon-2019-0650
49. Li F, Ma J, Yan C, Qi Y. ER stress-related mRNA-lncRNA co-expression gene signature predicts the prognosis and immune implications of esophageal cancer. *Am J Transl Res*. 2022;14(11):8064–8084. PMID:36505280. PMCID:PMC9730056.
50. Yi J, Wang L, Hu G, et al. CircPVT1 promotes ER-positive breast tumorigenesis and drug resistance by targeting *ESR1* and MAVS. *EMBO J*. 2023;42(10):e112408. doi:10.15252/embj.2022112408
51. Braga ACS, Carneiro BM, Batista MN, Akinaga MM, Bittar C, Rahal P. Heat shock proteins HSPB8 and DNAJC5B have HCV antiviral activity. *PLoS One*. 2017;12(11):e0188467. doi:10.1371/journal.pone.0188467
52. Sun H, Cai X, Zhou H, et al. The protein–protein interaction network and clinical significance of heat-shock proteins in esophageal squamous cell carcinoma. *Amino Acids*. 2018;50(6):685–697. doi:10.1007/s00726-018-2569-8
53. Zhou MH, Wang XK. Microenvironment-related prognostic genes in esophageal cancer. *Transl Cancer Res*. 2020;9(12):7531–7539. doi:10.21037/tcr-20-2288
54. Mirzaei MR, Asadi M, Mowla SJ, Hassanshahi G, Ahmadi Z. Down-regulation of *HSP40* gene family following OCT4B1 suppression in human tumor cell lines. *Iran J Basic Med Sci*. 2016;19(2):187–193. PMID:27081464. PMCID:PMC4818367.