

RESEARCH

Open Access



Endoplasmic reticulum stress-related gene model predicts prognosis and guides therapies in lung adenocarcinoma

Yuqi Song¹, Jianzun Ma¹, Linan Fang¹, Mingbo Tang¹, Xinliang Gao¹, Dongshan Zhu¹ and Wei Liu^{1*}

*Correspondence:
l_w01@jlu.edu.cn

¹ Department of Thoracic Surgery, The First Hospital of Jilin University, Changchun, Jilin, People's Republic of China

Abstract

Background: The prognosis and survival of lung adenocarcinoma (LUAD) patients are still not promising despite recent breakthroughs in treatment. Endoplasmic reticulum stress (ERS) is a self-protective mechanism resulting from an imbalance in quality control of unfolded proteins when cells are stressed, which plays an active role in lung cancer development, but the relationship between ERS and the pathological characteristics and clinical prognosis of LUAD patients remains unclear.

Methods: LASSO and Cox regression were applied based on sequencing information to construct the model, which was validated to be robust. The risk scores of the patients were calculated using the formula provided by the model, and the patients were divided into high and low-risk groups according to the median cut-off of risk scores. Cox regression analysis identifies independent prognostic factors for these patients, and enrichment analysis of prognosis-related genes was also performed. The relationship between risk scores and tumor mutation burden (TMB), cancer stem cell index, and drug sensitivity was explored.

Results: We constructed a 13-gene prognostic model for LUAD patients. Patients in the high-risk group had worse overall survival, lower immune score and ESTIMATE score, higher TMB, higher cancer stem cell index, and higher sensitivity to conventional chemotherapeutic agents. In addition, we constructed a nomogram that predicts 5-year survival in LUAD patients, which helps clinicians to foresee the prognosis from a new perspective.

Conclusions: Our results highlight the association of ERS with LUAD and the potential use of ERS in guiding treatment.

Keywords: Lung Adenocarcinoma, Endoplasmic Reticulum Stress, Gene Model, Prognosis, Therapy

Introduction

As one of the most prevalent types of malignant tumors, lung cancer has been ranked among the highest incidence and mortality rates in many parts of the world [1]. Approximately 85% of lung cancers are non-small cell lung cancers (NSCLC) [2]. As the most predominant subtype of NSCLC, lung adenocarcinoma (LUAD) accounts



© The Author(s) 2023. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

for approximately 50% of the total lung cancers [3]. Despite numerous breakthroughs in treating lung cancer, the 5-year survival rate after lung cancer is determined to be 10–20% in most nations [4]. Although targeted therapies have improved the prognosis of some patients, drug resistance inevitably occurs, so this group of patients has to choose other treatment modalities. Immunotherapy as a highly promising treatment has led to sustained remission and elongated survival in a certain proportion of patients, but 80% of progressive NSCLC patients do not have a response to the now confirmed immunotherapy with a single checkpoint inhibitor [5, 6]. Platinum-based chemotherapy stays the gold standard of treatment for advanced lung cancer patients [7].

Endoplasmic reticulum stress (ERS) is a state induced by increased synthesis of intracellular proteins, accumulation of mal-folded proteins, changes in cytosolic calcium levels, or imbalances in endoplasmic reticulum (ER) redox homeostasis [8]. However, if the stress on the ER is chronic or severe and quality control of protein folding cannot be met, ER activates its stress receptor, the unfolded protein response, which triggers programmed cell death [9]. This might be the pathogenesis of many diseases [10, 11], including type 2 diabetes, neurodegenerative diseases, and atherosclerosis. It was suggested that the ERS signaling pathway also plays an active part in cancer progression [12]. Serving as a self-protective mechanism, ERS is not only engaged in the proliferation of cancer cells in hypoxic environments [13], it also enables lung cancer cells to acquire resistance to chemotherapy, radiotherapy and targeted therapy [14]. Meanwhile, ERS as a double-edged sword is also closely related to drug-induced apoptosis of lung cancer cells [15–17].

A growing number of studies imply that ERS is tightly associated with the development of LUAD. However, the relationship between ERS and the pathological characteristics and clinical prognosis of LUAD patients is still unknown and needs to be further explored. At this stage, research on the use of specific biomarkers to construct prognostic models for lung cancer is emerging, and bioinformatics plays an important role during the process. Within this context, we constructed a prognostic model of ERS-related genes, combining the risk scores in the model with clinical characteristics (age, sex, staging, etc.) to establish a nomogram, and found its performance in estimating patient survival was excellent. Finally, we investigated the relationship between risk score and immune cell infiltration, cancer stem cell index, and drug sensitivity in the expectation of guiding individualized treatment of LUAD patients.

Materials and methods

Acquisition and processing of data

LUAD patients' RNA sequencing files and the corresponding clinicopathological profiles were downloaded from the TCGA database (<https://portal.gdc.cancer.gov/>) and the GSE72094 dataset in the GEO database (<https://www.ncbi.nlm.nih.gov/geo/>). Patients with missing overall survival data were excluded. Ultimately, 902 LUAD patients were included in the downstream analysis (504 in TCGA and 398 in the GSE72094 dataset). Fragments Per Kilobase of exon model per Million mapped fragments (FPKM) values were transformed into transcript volume per million (TPM) values in RNAseq transcriptome data, and R package limma [18] and sva [19] were used to perform batch correction and normalization of RNA-seq obtained from the two platforms. ERS-related

genes were downloaded from the GeneCards database (<https://www.genecards.org>), an integrated bioinformatics database that provides detailed information on all genes that are currently annotated and predictable in humans, and genes with relevance scores ≥ 10 were selected for this study. 376 ERS-related genes were obtained and collected in Additional file 1: Table S1.

After excluding data with incomplete information, we randomized altogether 902 patients in the TCGA database ($n = 504$) and the GSE72094 dataset of the GEO database ($n = 398$) as a TRAIN set ($n = 542$) and a TEST set ($n = 360$). TRAIN set was employed for model construction and the TEST set for validation. Expression data from all patients were used for nomogram construction, immune microenvironment analysis, and subsequent correlation analysis. GSE31210 ($n = 226$) from the GEO database was used for external validation of the prognostic model. The clinicopathological characteristics of all patients were summarized in Table 1.

Development of a prognostic model with ERS-related gene

The ERS-associated genes associated with overall survival were derived from a univariate Cox regression analysis of the combined standardized data from both datasets. Afterward, R package "glmnet"[20] was used to perform LASSO regression analysis on the prognostic data, and the penalty function was optimized using cross-validation. Prognosis-related genes with $p < 0.05$ were eventually included in the multifactorial Cox regression analysis to develop a prognostic prediction model for LUAD patients consisting of ERS-related genes. The formula is as follows:

$$RiskScore = \sum_{i=1}^n Coef_i * expressionvalue_i$$

where $Coef_i$ represents the coefficient of each prognostic gene and $expressionvalue_i$ means the relative expression level of each gene normalized by z-score. Subsequently, all those patients were split into two groups, high and low-risk, depending on the median risk score.

The risk score and independent prognostic analysis

The "survival" package in R was used to plot Kaplan–Meier survival curves, and Receiver Operating Characteristic (ROC) curves for each group were generated using the "timeROC" package [21] to show the relationship between the sensitivity and specificity of the model. The "pheatmap" package in R was utilized to display the risk score and survival status distribution for each patient, and the R package was also used to plot the prognostic gene expression heat maps.

Cox regression analyzed whether risk score and clinical characteristics were independent prognostic factors for patients with LUAD. The relationship between risk score models and the patient prognosis was validated in stratified analyses.

Construction and validation of nomogram

Using the "survival" [22] and "rms" [23] packages, a nomogram was constructed in terms of patients' gender, age, pathological stage, and risk score for predicting the overall

Table 1 Clinical characteristics of LUAD patients in each dataset

Clinical characteristics	Total	%
TCGA	504	100
Survival status		
Alive	321	63.69
Dead	183	36.31
Age		
< = 65 years old	238	47.22
> 65 years old	256	50.79
Unknown	10	1.99
Gender		
Male	234	46.43
Female	270	53.57
Stage		
I	270	53.57
II	119	23.61
III	81	16.07
IV	26	5.16
Unknown	8	1.59
GSE72094	398	100
Survival status		
Alive	285	71.61
Dead	113	28.39
Age		
< = 65 years old	118	29.65
> 65 years old	280	70.35
Gender		
Male	176	44.22
Female	222	55.78
Stage		
I	254	63.82
II	67	16.83
III	57	14.32
IV	15	3.77
Unknown	5	1.26
GSE31210	226	100
Survival status		
Alive	191	84.51
Dead	35	15.49
Age		
< = 65 years old	176	77.88
> 65 years old	50	22.12
Gender		
Male	105	46.46
Female	121	53.54
Stage		
IA	114	50.44
IB	54	23.89
II	58	25.67

survival of LUAD patients over 1, 3, and 5 years. The "survminer" and "timeROC" packages were utilized to plot ROC curves and calibration curves for checking the validity and reliability of the nomogram.

Protein–protein interaction (PPI) network

Using the STRING online platform (<https://STRING-db.org/>), a PPI network was constructed for 13 ERS-related genes in the model. Medium confidence (0.04) was selected as the minimum required interaction score. The line color signifies the type of interactive evidence. Active interaction sources come from all the options provided by the site, such as text mining, experiments, databases, co-expression, neighborhood, gene fusion, and co-occurrence.

Enrichment analysis

Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) were employed to annotate and functionally analyze prognosis-related genes using the "clusterProfiler" R package [24], with a filter of the adjusted p -value (q -value) < 0.05. GO enrichment analysis covers three aspects of biology: molecular function, cellular component, and biological processes.

Gene set variation analysis (GSVA) is a non-parametric, unsupervised method for assessing transcriptomic and genomic enrichment [25]. GSVA translates genetic level changes into pathway-level by performing a composite score on the set of genes of interest to determine the bio function of the specimen. In the study, pathway information was obtained from the Molecular Signatures Database (<https://www.gsea-msigdb.org/gsea/index.jsp>) (VERSION V7.0) and the GSVA algorithm was used to aggregate each gene set for scoring.

Tumor immune microenvironment analysis

Based on gene expression data, CIBERSORT can be used to assess the abundance of different cell types in mixed cell populations [26]. RNA-seq data from LUAD patients were analyzed using the CIBERSORT algorithm, and the relative proportions of 22 immune infiltrating cells were calculated. Spearman correlation analysis was used to explore the relationship between genes and immune cells.

The ESTIMATE algorithm calculates the stromal score, immune score, and ESTIMATE score of the tumor to reflect the purity of the tumor [27]. The "ggpubr" package was used to draw violin plots to graphically display the differences in each score.

Correlation analysis

The information on Tumor mutation burden (TMB) in the LUAD cohort was obtained from the TCGA database and analyzed with the "maftools" R package [28]. The number of somatic nonsynonymous point mutations is displayed in a waterfall plot for each individual. Differences in TMB between high and low-risk groups were also compared and presented as a box plot. In addition, cancer stem cell score calculation files were downloaded from the TCGA database, and the relationship between risk score and cancer stem cell index was evaluated.

Pharmaceutical sensitivity analysis

Differences in the half-inhibitory concentrations (IC50) of antitumor drugs in high and low-risk groups of LUAD patients were assessed using the "pRophetic" R package [29], according to the largest database of pharmacogenomics and presented in box plots.

Statistical analysis

Statistical analysis in this study was performed using the R programming language (version 4.1.1). Genes associated with prognosis and independent prognostic factors were identified using Cox regression analysis. The Kaplan–Meier curve and log-ranch test were applied to analyze differences in survival. In all analyses, *P* values < 0.05 were deemed to be statistically significant.

Results

The overall design and flow chart of this study is shown in Fig. 1.

Prognostic model calculates risk scores to help with patient grouping

By integrating the expression of ERS-related genes with patients’ clinical data, we conducted a univariate Cox regression analysis that yielded up to a sum of 96 ERS-associated genes significantly related to overall survival (Additional file 1: Table S2). Following that, LASSO regression analysis was performed in the TRAIN set, and the 21 genes (EIF2AK3, EIF2S1, DERL1, BCAP31, CAT, SLC6A4, G3BP1, HSPA4, ABL1, HSP90AA1, MBTPS2, PKP2, CKAP4, TXN, CASP9, VEGFA, CAV3, PRKCD, SERPINH1, DSP, SLC2A1) corresponding to the lambda.min values were selected and included in the multivariate Cox regression to construct a model including 13 ERS-related genes (Fig. 2A, B), with the hazard ratio (HR), *p*-value, and coefficient of the genes being shown in Fig. 2C. The risk score for LUAD patients can be computed with the following penalized function

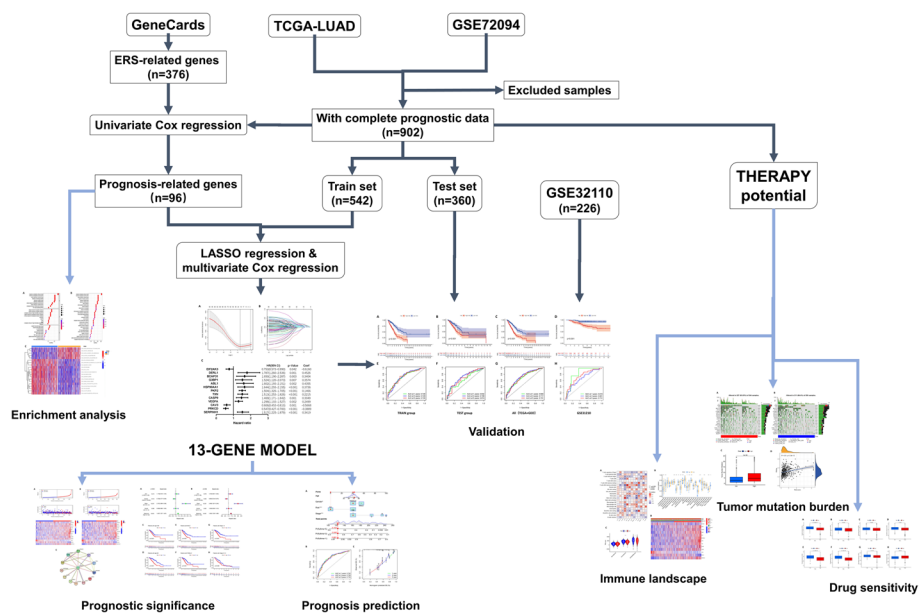


Fig. 1 Flowchart of this study

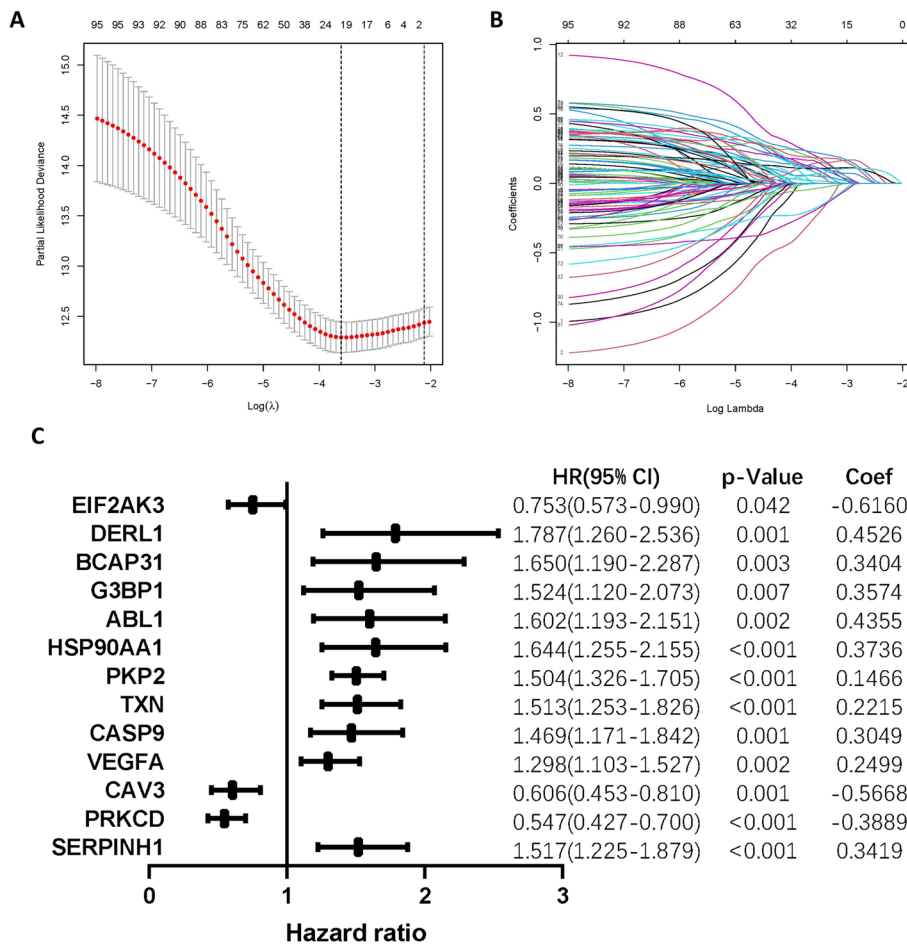


Fig. 2 Construction of a prognostic model. **A** The prognostic model is constructed by selecting the variable with the smallest partial likelihood of deviance. **B** tenfold cross-validation for screening variables in LASSO analysis. **C** ERS-related Genes and their correlation coefficients in the prognostic model

equation: Risk Score = $(- 0.6160 \times \text{EIF2AK3 expression}) + (0.4526 \times \text{DERL1 expression}) + (0.3404 \times \text{BCAP31 expression}) + (0.3574 \times \text{G3BP1 expression}) + (0.4355 \times \text{ABL1 expression}) + (0.3736 \times \text{HSP90AA1 expression}) + (0.1466 \times \text{PKP2 expression}) + (0.2215 \times \text{TXN expression}) + (0.3049 \times \text{CASP9 expression}) + (0.2499 \times \text{VEGFA expression}) + (- 0.5668 \times \text{CAV3 expression}) + (-0.3889 \times \text{PRKCD expression}) + (0.3419 \times \text{SERPINH1 expression})$.

We separated the entire sample into high and low-risk groups depending on the median risk score and plotted Kaplan–Meier survival curves. The curves showed that the survival rate and survival time were much lower in the high-risk group than in the low-risk group among both TRAIN and TEST sets (Fig. 3A–C, $p < 0.001$). Time-dependent ROC curves were applied to estimate the model (Fig. 3E–G), with area under the curve (AUC) values reaching 0.728 at 1, 3, and 5 years for the TRAIN set and 0.632, 0.706, and 0.674 at 1, 3, and 5 years for the TEST set. In addition, the GSE31210 dataset was used as an external validation, and the survival curves and ROC curves plotted on its basis as shown in Fig. 3D, H also helped us to validate the good performance of the model for prognosis assessment.

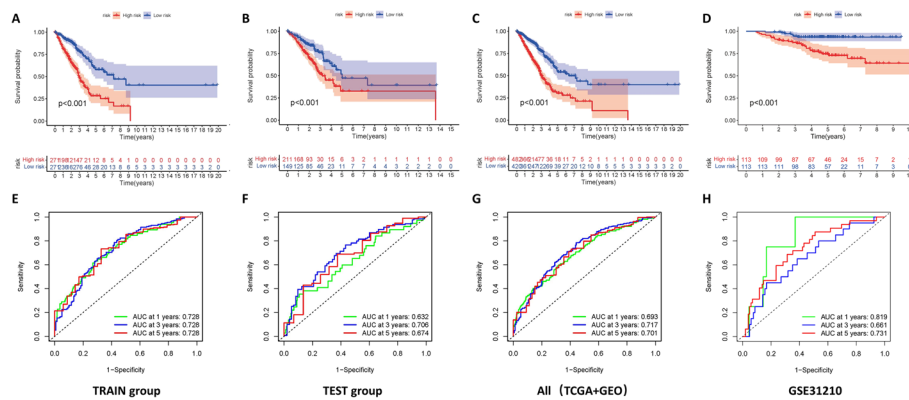


Fig. 3 A combination of internal and external validation verifies the predictive performance of the model. Kaplan-Meier curve analysis showed the difference in overall survival between the high and low-risk groups in the TRAIN set (A), the TEST set (B), and overall patients (C). D GSE31210 was used as an external validation dataset. E–H AUC was obtained from the ROC curve to verify the predictive efficacy of the overall survival time at 1, 3, and 5 years

The trends in risk scores, the distribution of survival status, and the relationship between the 13 genes included in the model and the risk groups are presented in Fig. 4A, B. The heat map suggests that high expression of DERL1, BCAP31, G3BP1, ABL1, HSP90AA1, PKP2, TXN, CASP9, VEGFA, and SERPINH1 was associated with high mortality in our high-risk group, while high expression of EIF2AK3, CAV3, and PRKCD was often found in the samples with low-risk scores.

The PPI networks of these 13 genes were mapped to give a better picture of the correlation between the proteins encoded by these genes (Fig. 4C).

Risk score can be an independent prognostic factor for LUAD patients

We performed univariate and multivariate Cox analyses to explore independent prognostic factors in patients with LUAD. Univariate Cox regression analysis revealed that risk scores (HR = 1.407. 95% confidence interval [CI] 1.337–1.481, $p < 0.001$) were significantly related to the prognosis of patients. In addition, gender (HR= 0.736. 95% CI 0.577–0.939, $p < 0.05$), and pathological stage (HR= 2.707. 95% CI 2.089–3.508, $p < 0.001$) were also observed as prognostic factors in patients with LUAD (Fig. 5A). After controlling for confounding variables, multivariate Cox regression analysis disclosed that the 13-gene risk score (HR = 1.405, 95% CI 1.329–1.486, $p < 0.001$) remained an indicator of patient overall survival, and similarly, gender and pathological stage also remained prognostic factors for these patients (Fig. 5B).

We further discussed the applicability of the risk score in the stratified analysis (Fig. 5C–H). Survival analysis suggested marked survival differences between the high and low-risk groups in terms of stratification characteristics including age, gender, and pathological stage. This also validated the robustness of the risk score’s ability to predict prognosis.

Nomogram accurately predict prognosis in LUAD patients

Based on all sample data obtained, we developed a nomogram including age, sex, stage, and risk score of the model to predict the survival time of patients (Fig. 6A). The

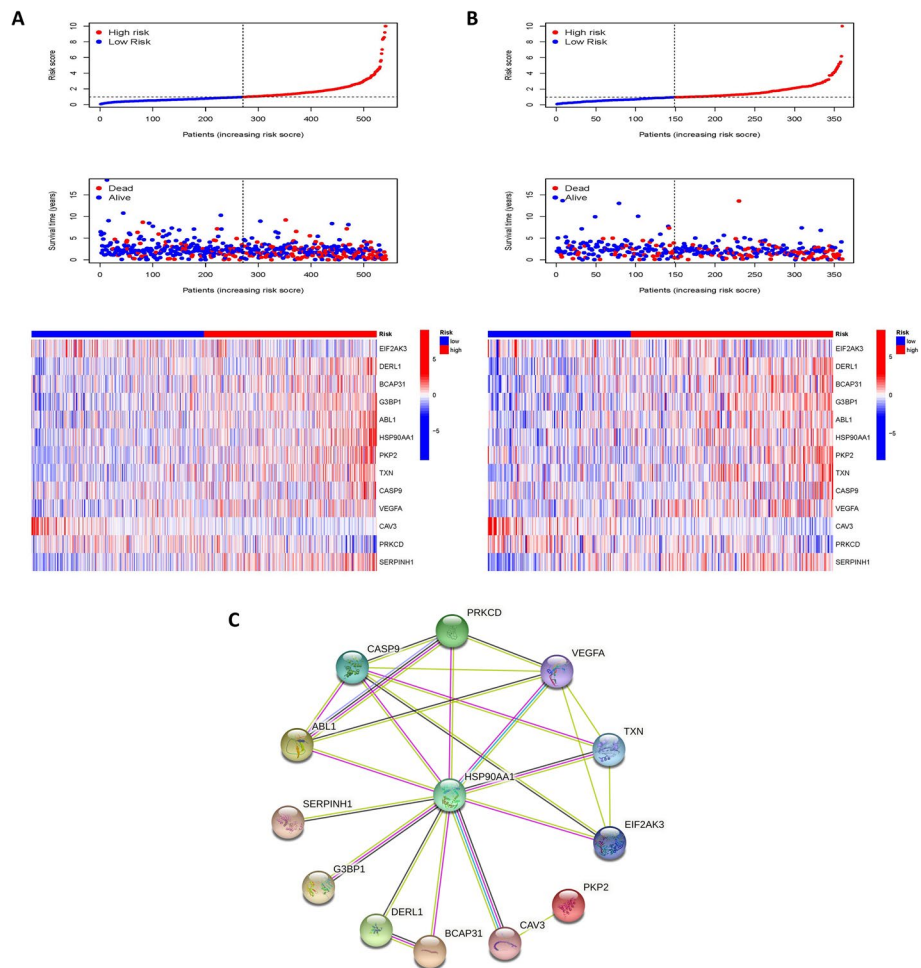


Fig. 4 Risk scores distinguish high and low-risk groups. Relationships between risk scores, distribution of survival status, and expression profiles of 13 ERS-related genes in the TRAIN set (A) and TEST set (B). C PPI network of the 13 genes in the model. The minimum required interaction score was set to medium confidence 0.400 and the line color indicates the type of interaction evidence

nomogram showed that the distribution of different clinical indicators and risk score values contribute differently to the overall scoring process, with the total score broadly forecasting a patient’s survival over 1, 3, and 5 years. The ROC curve was used to test the nomogram, and the results showed that the AUC was 0.742, 0.756, and 0.735 for 1, 3, and 5 years, respectively (Fig. 6B), which, combined with the calibration curve demonstrated in Fig. 6C, indicated that the nomogram had excellent predictive power for prognosis.

Identification of specific signals associated with prognosis

We found 96 prognostically relevant ERS-related genes significantly enriched in many pathways through GO and KEGG enrichment analysis. For example, a large number of genes were enriched in response to ERS, topologically incorrect protein, cellular response to chemical stress, etc. in the GO enrichment analysis (Fig. 7A). The

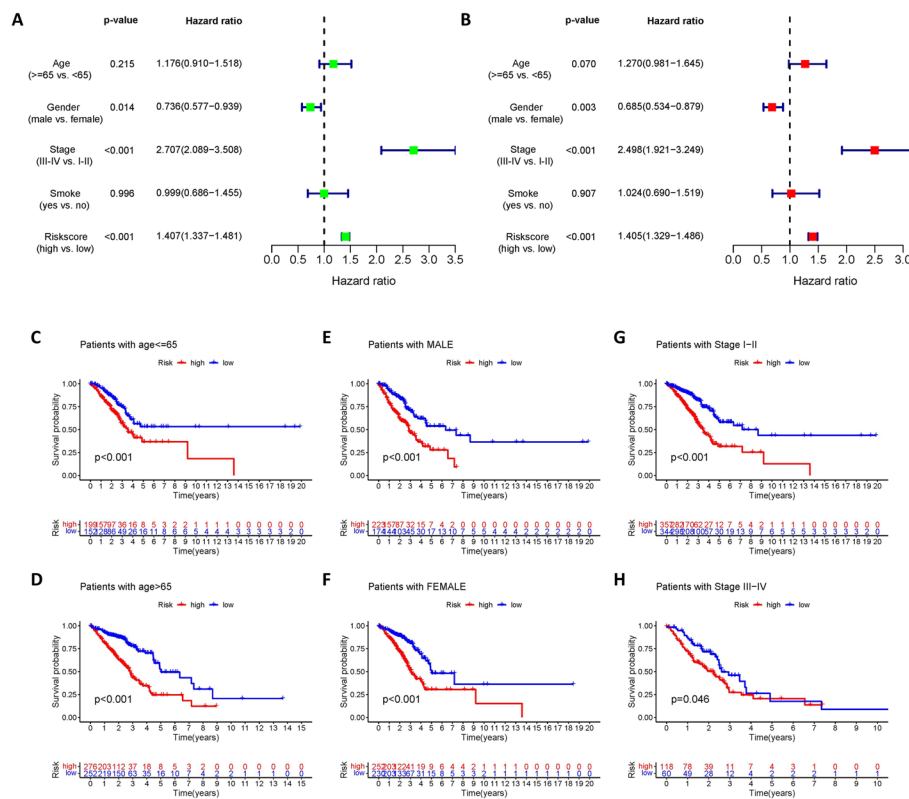


Fig. 5 Independent prognostic analysis and stratified analysis. Forest plot of univariate (A) and multivariate (B) Cox regression analysis based on the clinical information and risk score. Kaplan–Meier survival curves showed significant survival differences between the high and low-risk groups in a clinically stratified analysis including age (C–D), sex (E–F), and pathological stage (G–H)

enrichment of KEGG (Fig. 7B) displayed that key genes are primarily engaged in the pathways of response to oxidative stress, response to unfolded protein, and metal ion transport.

Next, we investigated the specific signaling pathways involved in the two risk groups using GSVA analysis and presented them in a heat map (Fig. 7C) and explored the potential molecular mechanisms affecting tumor progression. According to the GSVA results (Additional file 1: Table S3), the differences between the two groups focused on DNA REPLICATION, HOMOLOGOUS RECOMBINATION, CELL CYCLE, MISMATCH REPAIR, and PROTEASOME. These findings suggest that disturbances in the above-mentioned signaling pathways may have an impact on the prognosis of LUAD patients.

Significant differences in immune cell infiltration between 2 groups

It was shown that ERS has a crucial role in immune cells within the tumor microenvironment of LUAD [30]. Therefore, we further investigated whether ERS-related genes affect immune cells. The CIBERSORT algorithm was utilized to unearth the correlations between immune cells and genes in the model and we plotted a heat map to show them (Fig. 8A).

We then analyzed the difference in the content of immune cells in high and low-risk groups. Figure 8B showed that the content of immune cells such as activated B cells,

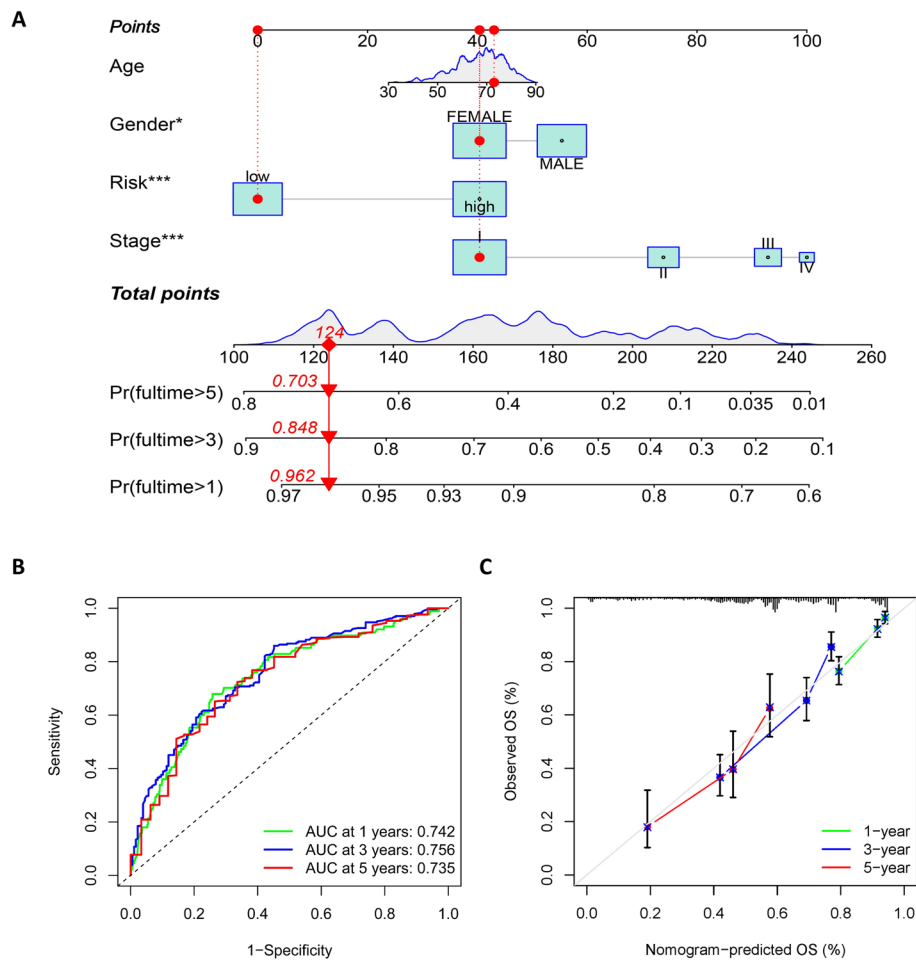


Fig. 6 Nomogram is a good tool for predicting prognosis. **A** A nomogram was developed including patient's age, gender, stage, and risk score to predict survival at 1,3,5 year. ROC curves (**B**) and calibration curves (**C**) show that the nomogram has good predictive performance

immature B cells, and immature dendritic cells was noticeably lower in the high-risk group, whereas, activated CD4 T cells, CD56dim natural killer cells, and type2 T helper cells were clearly higher.

The tumor microenvironment consists of stromal cells and infiltrating immune cells, reflecting the purity of the tumor. As the number of stromal cells and infiltrating immune cells increases, tumor purity decreases. The results of the ESTIMATE analysis displayed that immune scores and ESTIMATE scores were markedly lower in the high-risk group (Fig. 8C). We further plotted the heat map (Fig. 8D) with the information on age, gender, etc. to better demonstrate the relationship between clinical phenotype, key genes involved in the model, and high and low-risk groups.

Greater TMB and higher CSC index in high-risk group

Tumor mutation burden (TMB) is simply the total number of mutations carried by tumor cells, specifically the number of somatic mutations per megabyte of the genomic sequence being interrogated. Patients with high TMB have been reported to have higher clinical benefits after immunotherapy, and thus TMB can be considered a potential

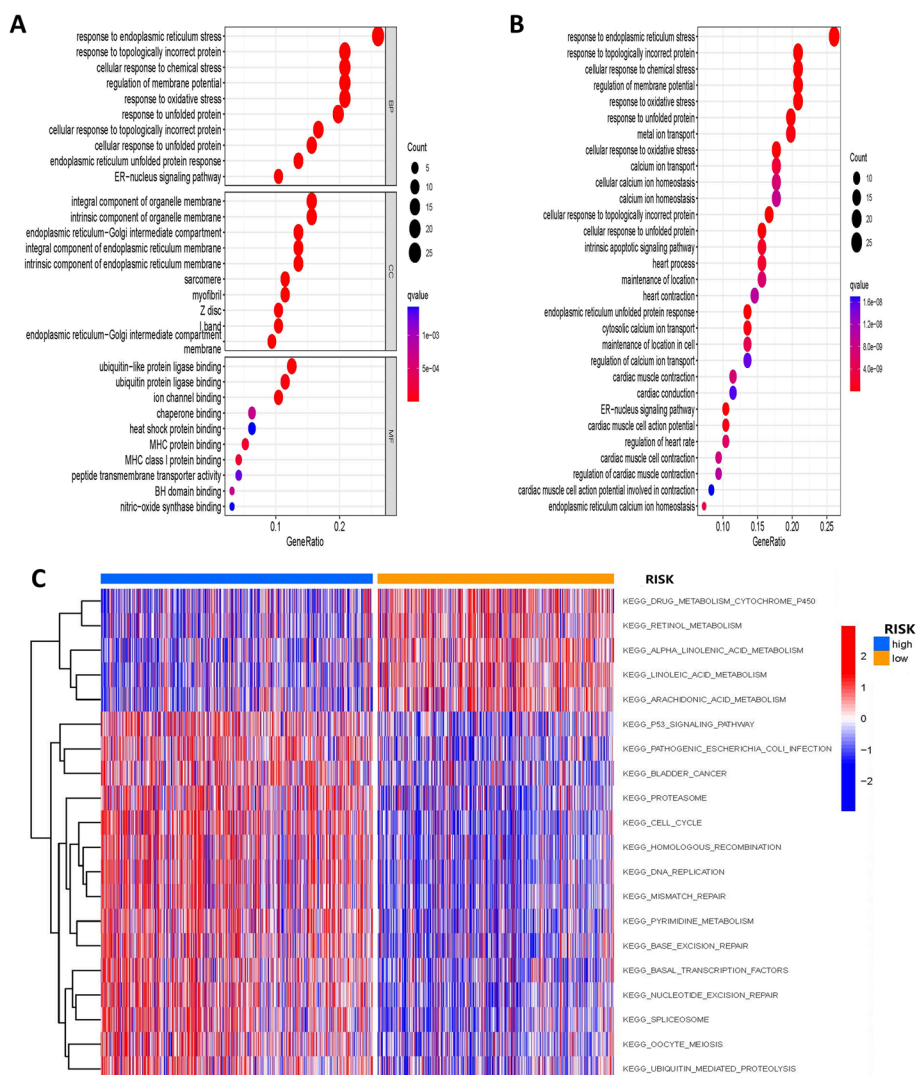


Fig. 7 Enrichment analysis of prognostic genes. GO (A) and KEGG (B) pathway analysis of 96 prognostically relevant ERS-related genes. C GSEA analysis reveals differences between specific signaling pathways involved in the two risk groups

molecular diagnostic marker for tumor immune checkpoint inhibitors to guide patients receiving treatment [31].

Mutation frequencies for the high and low-risk groups were shown in the waterfall plot Fig. 9A, B, which showed that there is no difference in the names of the top mutated genes in the high and low-risk groups, but the number of mutations differs. We then performed a correlation analysis of TMB (Fig. 9C) and there was a significant difference in TMB between the two groups.

Cancer stem cells (CSCs) are a subpopulation of cells in tumors with the ability to self-renew, which are highly correlated with tumor development and play an important role in other malignant phenotypes such as cancer metastasis and recurrence [32]. Maciej Wiznerowicz et al. generated a cancer stem cell index (mRNasi) derived from the mRNA expression of 512 LUAD patients using a one-class logistic regression

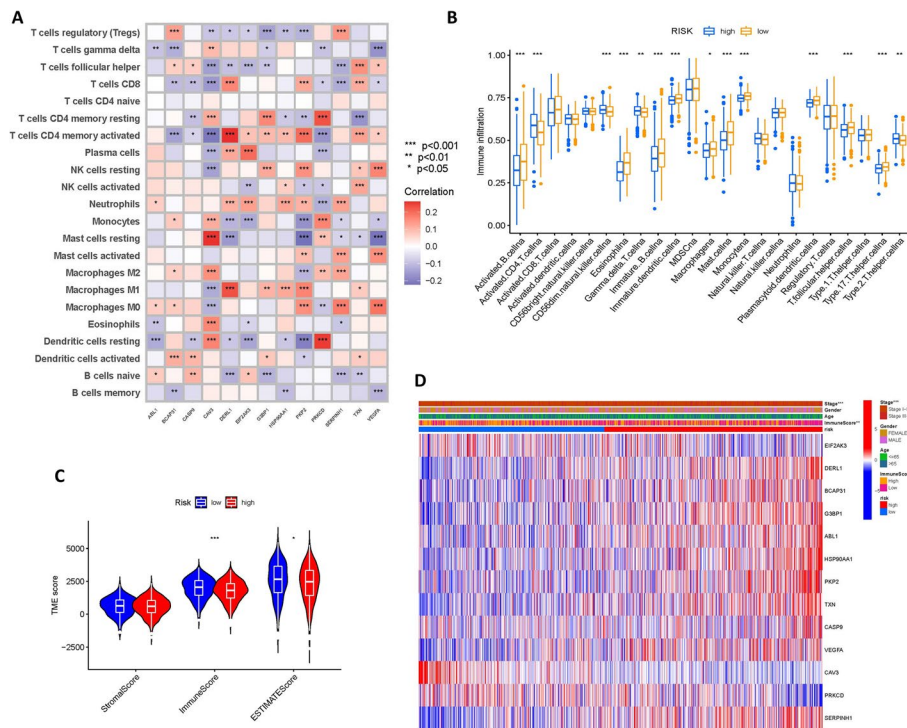


Fig. 8 Landscape of immune cell infiltration. **A** Correlation heat map between immune cells and genes in the model. **B** Differences in immune cells between high and low-risk groups. **C** ESTIMATE algorithm was applied to predict the tumor purity and the presence of infiltrating immune/stromal cells in tumor tissues. **D** Heat map of correlations between clinical features, risk score, immune score groups, and genes in the model

machine-learning algorithm (OCLR) [33]. Spearman analysis reveals a positive correlation between mRNasi and risk score, with a correlation coefficient of 0.31, $p < 0.001$ (Fig. 9D).

Predicted differences in drug treatment effects

Chemotherapy, as a classical lung cancer treatment, is widely used in the adjuvant and neoadjuvant treatment of NSCLC [34]. In this study, the "pRRophetic" R package was used to predict the IC50 of antitumor drugs for each tumor specimen and to further investigate the difference in sensitivity between the two groups. The lower the IC50 value, the more sensitive the drug. From Fig. 10A–G we can see that the common chemotherapeutic agents (cisplatin, paclitaxel, gemcitabine, vinorelbine, docetaxel, doxorubicin, etoposide) are all more sensitive in the high-risk group, while for the erlotinib (Fig. 10H), the sensitivity is significantly higher in the low-risk group, implying that the low-risk group may benefit more from EGFR-TKI therapy.

Discussion

The endoplasmic reticulum (ER) is an organelle that is essential to all eukaryotic cells and is involved in vital cellular activities. Factors such as hypoxia, hypoglycemia, high temperature, acidosis, calcium levels, redox environment, and energy levels can affect and interfere with the normal function of the ER, leading to ERS and affecting protein folding in the ER lumen [35]. There is growing evidence showing that ERS is

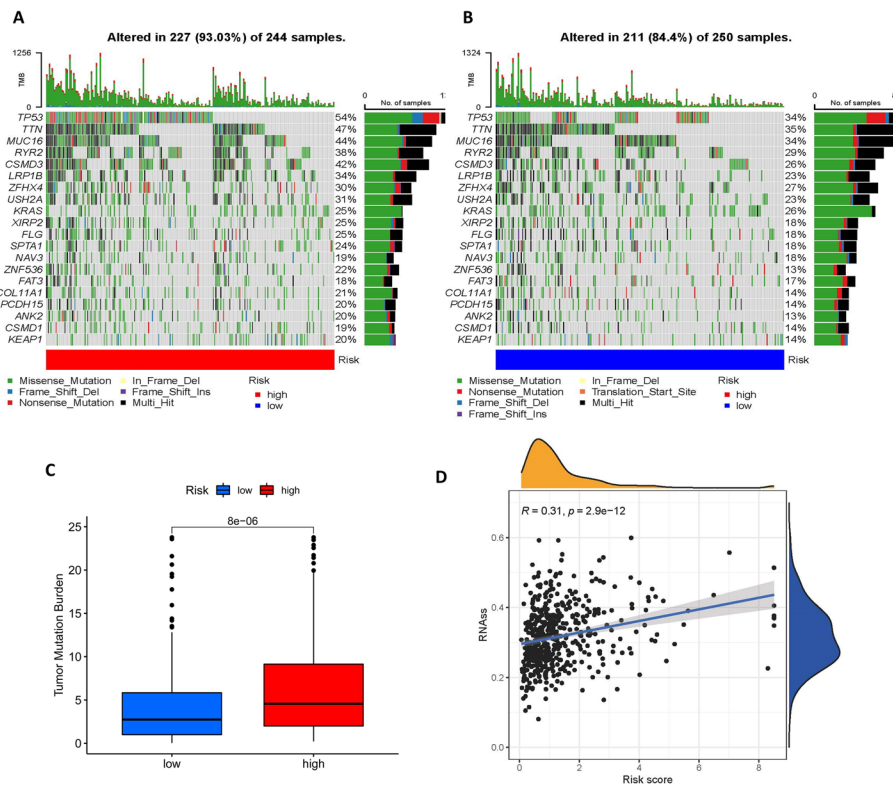


Fig. 9 Oncogenic Mutation Landscape and cancer stem cell index. **A–B** Waterfall plot of the frequency of mutations in high and low-risk groups. **C** Statistically significant differences in TMB between high and low-risk groups. **D** Risk score is positively correlated with cancer stem cell index

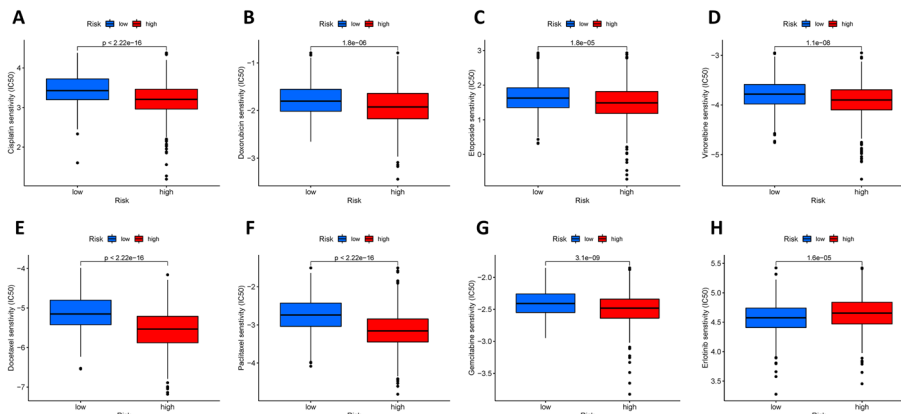


Fig. 10 Drug sensitivity analysis. **A–G** The high-risk group is more sensitive to common chemotherapy drugs (cisplatin, paclitaxel, gemcitabine, vinorelbine, docetaxel, doxorubicin, etoposide). **H** The low-risk group has higher sensitivity to the EGFR-TKI erlotinib

an important pathway for cancer cell survival. Due to some characteristics of cancer cells, there is an increased need for protein processing in the ER and the mutant proteins expressed by cancer cells do not fold properly, thereby activating the ERS. The high metabolism of the tumor mass leads to a deterioration of the growth environment such as nutritional deficiencies, and these processes are also potent triggers

for the initiation of ERS, resulting in greater tumorigenicity, metastasis, and drug resistance in tumor cells [36]. Interestingly, however, there is sufficient evidence in the literature that the use of inducers to enhance ERS is an effective strategy to induce apoptosis in cancer cells [37-39].

For lung cancer, ERS can induce cellular autophagy in LUAD cells and promote the survival of drug-treated tumors, accelerating drug resistance, recurrence, and malignant development of tumors [40]. In turn, the antitumor activity of some other compounds is achieved through apoptosis mediated with ERS [41, 42], implying that precise modulation of ERS can lead to tumor elimination. Notably, the relationship between ERS and LUAD is not very clear, in which the tumor immune micro-environment may play an important role [43]. So, a thorough comprehension of the mechanisms and results in this field is essential to translate our knowledge into new therapeutic approaches.

In the present study, we applied LASSO regression to construct a 13-gene prognostic model for LUAD patients. Among these genes, high expression of DERL1, BCAP31, G3BP1, ABL1, HSP90AA1, PKP2, TXN, CASP9, VEGFA, and SERPINH1 was highly positively correlated with poorer prognosis. HSP90AA1 (Heat Shock Protein 90 Alpha Family Class A Member 1), as an important node in the PPI network, is a hot topic in the pathogenesis of malignant tumors. Studies have shown that the product of the HSP90AA1 protein, HSP90 α , is known to function in the regulation of tumor invasion and migration [44]. In addition, high expression of HSP90AA1 in lung cancer cells is closely associated with lung cancer progression and treatment response [45]. VEGFA (Vascular Endothelial Growth Factor A) is an important regulator of tumor angiogenesis in many solid tumors [46]. Clinical advances in anti-VEGF angiogenesis therapy have been successful and have been translated and utilized in lung cancer [47]. SERPINH1 (Serpin Family H Member 1), also known as HSP47, is an important stress-related protein on the endoplasmic reticulum [48], with higher expression in squamous carcinoma tissues than in normal human bronchial epithelial cells [49]. Similarly, PKP2 (Plakophilin 2) and DERL1 (Derlin 1) are overexpressed in lung cancer cells [50, 51], promote the development of lung cancer, and are associated with lung cancer prognosis [52]. Overexpression of ABL1 (ABL Proto-Oncogene 1) gene is associated with shorter survival in LUAD patients, and animal studies suggest that its inhibitors are effective in treating metastatic NSCLC [53]. G3BP1 (G3BP Stress Granule Assembly Factor 1), BCAP31 (B Cell Receptor Associated Protein 31), CASP9 (Caspase 9), and TXN (Thioredoxin) are involved in multiple biological processes in tumors by regulating signaling pathways [54], thereby promoting tumor cell proliferation and metastasis as well as regulating apoptosis [55]. All of them may be potential targets for tumor therapy [56] and provide new strategies for the treatment of LUAD [57].

For the remaining three genes with Coef less than 0 in the risk score, EIF2AK3 (Eukaryotic Translation Initiation Factor 2 Alpha Kinase 3), which encodes a membrane protein, silencing in cancer cells prevents ERS and induces apoptosis [58]. PRKCD (Protein Kinase C Delta) genes play a key role in growth inhibition, differentiation, apoptosis, and tumor suppression [59, 60]. CAV3 (Caveolin 3) mainly encodes a protein of caveolin-3, which is mainly distributed in the membranes surrounding

myocytes, and its dysfunction is associated with many diseases, including cancer and diabetes [61, 62]. Notably, the interrelationship of some genes with LUAD has not received much attention from researchers previously. These genes were identified and given significance in our study and remain to be further investigated in the future.

In the tumor microenvironment, ERS can serve as a novel biological marker [63]. Tumor cells themselves are in a variety of metabolic abnormalities that disrupt the homeostasis of intracellular protein folding and thus induce ERS in cells. This simultaneously induces changes in the microenvironment, such as nutrient deficiency, hypoxia, and imbalance of redox reactions, triggering the occurrence of ERS within the infiltrating immune cells and affecting tumor immunity. For example, ERS affects the myeloid cells in suppressing antitumor immune surveillance [30], and ERS in cancer cells influences the recruitment and function of immune cells, modulates T cell-mediated immune responses, and NK cell-mediated tumor recognition [63], but the exact mechanisms are not clear. Our results showed that the content of multiple immune cells was significantly different in high and low-risk groups, which may be due to disparities in the tumor microenvironment. Moreover, there is also a strong correlation between the genes included in the model and multiple immune cells, which provides clues to understanding the mechanism of ERS involvement in tumor immunity in the LUAD microenvironment. In addition, the TMB of LUAD patients increases with increasing risk scores, and studies have shown that patients with high TMB possess higher survival rates after receiving immunotherapy [64]. This makes the application of our model to immunotherapy possible, and its in-depth exploration will help guide the subsequent eradication of the tumor itself and improve the efficacy of immunotherapy.

CSCs are intimately tied to tumor heterogeneity, antitumor resistance, and metastasis of lung cancer. A significant negative correlation between CSCs and antitumor immunity has been observed [65]. In our results, the cancer stemness index was higher in the high-risk group, but its immune score was lower. This implied the possible existence of genes or pathways that maintain the stemness of tumor cells and inhibit the function of immune cells, thus regulating tumor progression [66]. Importantly, we should be aware that future research on CSC-targeted tumor therapy should be strengthened to enrich the means of tumor treatment [67].

To provide more accurate treatment options for LUAD patients, we used the R package "pRRophetic" to predict the differences in IC50 values of common antitumor drugs between the two groups on the basis of drug sensitivity data. According to the results, chemotherapy, one of the most important treatment modalities for LUAD, continues to benefit patients in the high-risk group (all common chemotherapeutic agents have a lower IC50), while the low-risk group has more treatment options, including targeted therapies (lower IC50 of Erlotinib). Unfortunately, information on immune checkpoint inhibitors is missing here, and our approach was not successful in predicting the sensitivity to immunotherapy in high and low-risk groups.

Of course, our study has some shortcomings. First, due to the limited data on LUAD patients in the public database, the sample size covered in this study was relatively insufficient, resulting in some results not representing the whole picture. Second, the dataset obtained from the public database lacks some important information, which limits us to more in-depth analysis. Furthermore, the 13 genes obtained in the prognosis-related

model also need to be validated externally at various levels (genes, proteins, in vivo experiments) to explore and reveal their specific mechanisms.

In summary, we constructed an ERS-related prognostic model and performed drug sensitivity analysis. Our results provide important clues to study the role of ERS in LUAD and are expected to help clinicians understand the relationship between ERS and patient prognosis from a novel perspective. We also expect that it can guide more appropriate drug treatment options for patients in the future.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12859-023-05384-z>

Additional file 1. Supplementary tables for details about ERS-related genes and GSVA analysis results.

Acknowledgements

The TCGA and GEO database is gratefully acknowledged for the availability of the raw research data.

Author contributions

Conceptualization: YS, LF, DZ; Methodology: YS, JM; Formal analysis and investigation: YS, JM; Writing—original draft preparation: YS, LF; Writing—review and editing: LF, MT, XG, DZ, WL; Resources: All authors; Supervision: WL. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (#82003284 to Linan Fang) and the Scientific Research Foundation of The First Hospital of Jilin University (#JDYY11202015 to Linan Fang).

Availability of data and materials

The datasets analyzed in the current study are available from the TCGA database [<https://portal.gdc.cancer.gov/>] and GEO database [<https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE72094>] repository.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 31 January 2023 Accepted: 13 June 2023

Published online: 16 June 2023

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, et al. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin.* 2021;71(3):209–49.
2. Molina JR, Yang P, Cassivi SD, Schild SE, Adjei AA. Non-small cell lung cancer: epidemiology, risk factors, treatment, and survivorship. *Mayo Clin Proc.* 2008;83(5):584–94.
3. Barta JA, Powell CA, Wisnivesky JP. Global epidemiology of lung cancer. *Ann Glob Health.* 2019. <https://doi.org/10.5334/aogh.2419>.
4. Allemani C, Matsuda T, Di Carlo V, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *Lancet.* 2018;391(10125):1023–75.
5. Brahmer JR, Tykodi SS, Chow LQ, Hwu WJ, Topalian SL, Hwu P, et al. Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N Engl J Med.* 2012;366(26):2455–65.
6. Topalian SL, Hodi FS, Brahmer JR, Gettinger SN, Smith DC, McDermott DF, et al. Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med.* 2012;366(26):2443–54.
7. Hanna N, Johnson D, Temin S, Baker S Jr, Brahmer J, Ellis PM, et al. Systemic therapy for stage IV non-small-cell lung cancer: American society of clinical oncology clinical practice guideline update. *J Clin Oncol.* 2017;35(30):3484–515.
8. Ron D, Walter P. Signal integration in the endoplasmic reticulum unfolded protein response. *Nat Rev Mol Cell Biol.* 2007;8(7):519–29.
9. Ogen-Shtern N, Ben David T, Lederkremer GZ. Protein aggregation and ER stress. *Brain Res.* 2016;1648(Pt B):658–66.
10. Tabas I, Ron D. Integrating the mechanisms of apoptosis induced by endoplasmic reticulum stress. *Nat Cell Biol.* 2011;13(3):184–90.

11. Marciniak SJ, Ron D. Endoplasmic reticulum stress signaling in disease. *Physiol Rev.* 2006;86(4):1133–49.
12. da Silva DC, Valentão P, Andrade PB, Pereira DM. Endoplasmic reticulum stress signaling in cancer and neurodegenerative disorders: Tools and strategies to understand its complexity. *Pharmacol Res.* 2020;155: 104702.
13. Yao X, Tu Y, Xu Y, Guo Y, Yao F, Zhang X. Endoplasmic reticulum stress confers 5-fluorouracil resistance in breast cancer cell via the GRP78/OCT4/lncRNA MIAT/AKT pathway. *Am J Cancer Res.* 2020;10(3):838–55.
14. Liao CH, Tzeng YT, Lai GM, Chang CL, Hu MH, Tsai WL, et al. Omega-3 fatty acid-enriched fish oil and selenium combination modulates endoplasmic reticulum stress response elements and reverses acquired gefitinib resistance in HCC827 lung adenocarcinoma cells. *Mar Drugs.* 2020;18(8):399.
15. Joo JH, Liao G, Collins JB, Grissom SF, Jetten AM. Farnesol-induced apoptosis in human lung carcinoma cells is coupled to the endoplasmic reticulum stress response. *Cancer Res.* 2007;67(16):7929–36.
16. Lu J, Li Y, Gong S, Wang J, Lu X, Jin Q, et al. Ciclopirox targets cellular bioenergetics and activates ER stress to induce apoptosis in non-small cell lung cancer cells. *Cell Commun Signal.* 2022;20(1):37.
17. Cheng WC, Wen Y, Chiu YS, Chou CH, Lim CJ, Lin SH, et al. Pendulone induces apoptosis via the ROS-mediated ER-stress pathway in human non-small cell lung cancer cells. *Toxicol In Vitro.* 2022;81: 105346.
18. Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W, et al. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* 2015;43(7): e47.
19. Leek JT, Johnson WE, Parker HS, Jaffe AE, Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics.* 2012;28(6):882–3.
20. Engebretsen S, Bohlin J. Statistical predictions with glmnet. *Clin Epigenetics.* 2019;11(1):123.
21. Blanche P, Dartigues JF, Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med.* 2013;32(30):5381–97.
22. Li JCA. Modeling survival data: extending the cox model. *Sociol Methods Res.* 2003;32(1):117–20.
23. Harrell FE Jr, Lee KL, Califf RM, Pryor DB, Rosati RA. Regression modelling strategies for improved prognostic prediction. *Stat Med.* 1984;3(2):143–52.
24. Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *OmicS.* 2012;16(5):284–7.
25. Hänzelmann S, Castelo R, Guinney J. GSEA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinf.* 2013;14:7.
26. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods.* 2015;12(5):453–7.
27. Chakraborty H, Hossain A. R package to estimate intracluster correlation coefficient with confidence interval for binary data. *Comput Methods Programs Biomed.* 2018;155:85–92.
28. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res.* 2018;28(11):1747–56.
29. Geeleher P, Cox N, Huang RS. pRRophetic: an R package for prediction of clinical chemotherapeutic response from tumor gene expression levels. *PLoS One.* 2014;9(9): e107468.
30. Mohamed E, Cao Y, Rodriguez PC. Endoplasmic reticulum stress regulates tumor growth and anti-tumor immunity: a promising opportunity for cancer immunotherapy. *Cancer Immunol Immunother.* 2017;66(8):1069–78.
31. Chan TA, Yarchoan M, Jaffee E, Swanton C, Quezada SA, Stenzinger A, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol.* 2019;30(1):44–56.
32. Ajani JA, Song S, Hochster HS, Steinberg IB. Cancer stem cells: the promise and the potential. *Semin Oncol.* 2015;42(Suppl 1):S3–17.
33. 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):338–54.e15.
34. Brahmer JR, Ettinger DS. Non-small cell lung cancer: adjuvant and neo-adjuvant chemotherapy. *Respirology.* 2007;12(3):320–5.
35. Malhotra JD, Kaufman RJ. The endoplasmic reticulum and the unfolded protein response. *Semin Cell Dev Biol.* 2007;18(6):716–31.
36. Cubillos-Ruiz JR, Bettigole SE, Glimcher LH. Tumorigenic and immunosuppressive effects of endoplasmic reticulum stress in cancer. *Cell.* 2017;168(4):692–706.
37. Ribeiro V, Andrade PB, Valentão P, Pereira DM. Benzoquinones from *Cyperus* spp trigger IRE1 α -independent and PERK-dependent ER stress in human stomach cancer cells and are novel proteasome inhibitors. *Phytomedicine.* 2019;63:153017.
38. Fan C, Yang Y, Liu Y, Jiang S, Di S, Hu W, et al. Icaritin displays anticancer activity against human esophageal cancer cells via regulating endoplasmic reticulum stress-mediated apoptotic signaling. *Sci Rep.* 2016;6:21145.
39. Cerezo M, Rocchi S. New anti-cancer molecules targeting HSPA5/BIP to induce endoplasmic reticulum stress, autophagy and apoptosis. *Autophagy.* 2017;13(1):216–7.
40. Wang H, Zhang G. Endoplasmic reticulum stress-mediated autophagy protects against β , β -dimethylacrylyshikonin-induced apoptosis in lung adenocarcinoma cells. *Cancer Sci.* 2018;109(6):1889–901.
41. Wu SH, Hang LW, Yang JS, Chen HY, Lin HY, Chiang JH, et al. Curcumin induces apoptosis in human non-small cell lung cancer NCI-H460 cells through ER stress and caspase cascade- and mitochondria-dependent pathways. *Anti-cancer Res.* 2010;30(6):2125–33.
42. Zhu J, Xu S, Gao W, Feng J, Zhao G. Honokiol induces endoplasmic reticulum stress-mediated apoptosis in human lung cancer cells. *Life Sci.* 2019;221:204–11.
43. Xia S, Duan W, Liu W, Zhang X, Wang Q. GRP78 in lung cancer. *J Transl Med.* 2021;19(1):118.
44. Wong DS, Jay DG. Emerging Roles of Extracellular Hsp90 in Cancer. *Adv Cancer Res.* 2016;129:141–63.
45. Shi Y, Liu X, Lou J, Han X, Zhang L, Wang Q, et al. Plasma levels of heat shock protein 90 alpha associated with lung cancer development and treatment responses. *Clin Cancer Res.* 2014;20(23):6016–22.
46. Dvorak HF. Vascular permeability factor/vascular endothelial growth factor: a critical cytokine in tumor angiogenesis and a potential target for diagnosis and therapy. *J Clin Oncol.* 2002;20(21):4368–80.
47. Ferrara N, Kerbel RS. Angiogenesis as a therapeutic target. *Nature.* 2005;438(7070):967–74.

48. Ito S, Nagata K. Biology of Hsp47 (Serp1H1), a collagen-specific molecular chaperone. *Semin Cell Dev Biol.* 2017;62:142–51.
49. Poschmann G, Sitek B, Sipos B, Ulrich A, Wiese S, Stephan C, et al. Identification of proteomic differences between squamous cell carcinoma of the lung and bronchial epithelium. *Mol Cell Proteomics.* 2009;8(5):1105–16.
50. Wu Y, Liu L, Shen X, Liu W, Ma R. Plakophilin-2 promotes lung adenocarcinoma development via enhancing focal adhesion and epithelial-mesenchymal transition. *Cancer Manag Res.* 2021;13:559–70.
51. Hao XL, Tian Z, Han F, Chen JP, Gao LY, Liu JY. Plakophilin-2 accelerates cell proliferation and migration through activating EGFR signaling in lung adenocarcinoma. *Pathol Res Pract.* 2019;215(7): 152438.
52. Mao M, Zhang J, Jiang J. Overexpression of derlin-1 is associated with poor prognosis in patients with non-small cell lung cancer. *Ann Clin Lab Sci.* 2018;48(1):29–34.
53. Gu JJ, Rouse C, Xu X, Wang J, Onaitis MW, Pendergast AM. Inactivation of ABL kinases suppresses non-small cell lung cancer metastasis. *JCI Insight.* 2016;1(21): e89647.
54. Ribback S, Winter S, Klatte T, Schaeffeler E, Gellert M, Stühler V, et al. Thioredoxin 1 (Trx1) is associated with poor prognosis in clear cell renal cell carcinoma (ccRCC): an example for the crucial role of redox signaling in ccRCC. *World J Urol.* 2022;40(3):739–46.
55. Wu GS, Ding Z. Caspase 9 is required for p53-dependent apoptosis and chemosensitivity in a human ovarian cancer cell line. *Oncogene.* 2002;21(1):1–8.
56. Zhang CH, Wang JX, Cai ML, Shao R, Liu H, Zhao WL. The roles and mechanisms of G3BP1 in tumour promotion. *J Drug Target.* 2019;27(3):300–5.
57. Wang J, Jiang D, Li Z, Yang S, Zhou J, Zhang G, et al. BCAP31, a cancer/testis antigen-like protein, can act as a probe for non-small-cell lung cancer metastasis. *Sci Rep.* 2020;10(1):4025.
58. Harding HP, Zhang Y, Ron D. Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. *Nature.* 1999;397(6716):271–4.
59. Brodie C, Blumberg PM. Regulation of cell apoptosis by protein kinase c delta. *Apoptosis.* 2003;8(1):19–27.
60. Jackson DN, Foster DA. The enigmatic protein kinase Cdelta: complex roles in cell proliferation and survival. *Faseb J.* 2004;18(6):627–36.
61. Stahlhut M, Sandvig K, van Deurs B. Caveolae: uniform structures with multiple functions in signaling, cell growth, and cancer. *Exp Cell Res.* 2000;261(1):111–8.
62. Pradhan BS, Prószyński TJ. A role for caveolin-3 in the pathogenesis of muscular dystrophies. *Int J Mol Sci.* 2020. <https://doi.org/10.3390/ijms21228736>.
63. Chen X, Cubillos-Ruiz JR. Endoplasmic reticulum stress signals in the tumour and its microenvironment. *Nat Rev Cancer.* 2021;21(2):71–88.
64. Cao D, Xu H, Xu X, Guo T, Ge W. High tumor mutation burden predicts better efficacy of immunotherapy: a pooled analysis of 103078 cancer patients. *Oncoimmunology.* 2019;8(9): e1629258.
65. Leon G, MacDonagh L, Finn SP, Cuffe S, Barr MP. Cancer stem cells in drug resistant lung cancer: targeting cell surface markers and signaling pathways. *Pharmacol Ther.* 2016;158:71–90.
66. Paczulla AM, Rothfelder K, Raffel S, Konantz M, Steinbacher J, Wang H, et al. Absence of NKG2D ligands defines leukaemia stem cells and mediates their immune evasion. *Nature.* 2019;572(7768):254–9.
67. Dianat-Moghadam H, Mahari A, Salahlou R, Khalili M, Azizi M, Sadeghzadeh H. Immune evader cancer stem cells direct the perspective approaches to cancer immunotherapy. *Stem Cell Res Ther.* 2022;13(1):150.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

