

RAPID COMMUNICATION

Construction and validation of a novel SUMOylation-related lncRNAs signature for predicting the prognosis, tumor immune microenvironment, and therapeutic sensitivity of lung adenocarcinoma

Lung adenocarcinoma (LUAD) is the most prevalent subtype of lung cancer and the leading cause of cancer deaths worldwide.¹ Despite immunotherapy, radiotherapy, and chemotherapy advances, treatment outcomes remain unsatisfactory.² Thus, prognostic models that accurately predict patient prognosis and guide individualized treatment are desperately needed. SUMOylation, a reversible post-translational modification, is a crucial molecular regulatory mechanism that affects tumor progression.³ However, the role of sumoylation-related lncRNAs (SR-lncRNAs) in LUAD has not been explored. This study aimed to construct and validate an SR-lncRNAs signature for predicting LUAD prognosis, tumor immune microenvironment, and therapeutic sensitivity.

Figure S1 shows the study's flowchart. Firstly, after removing genes expressed in less than half of patients, the TCGA_LUAD dataset yielded 15,831 lncRNAs. Then, 2858 lncRNAs were identified as differentially expressed lncRNAs based on false discovery rate < 0.01 and $|\log_2$ fold change| ≥ 1.0 (Fig. S2). We found 187 sumoylation-related genes in the Molecular Signatures Database (<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>) and collected expression data from the TCGA_LUAD dataset. Pearson's correlation analyses between differentially expressed lncRNAs and sumoylation-related genes yielded 831 SR-lncRNAs. Univariate Cox regression analyses further identified 54 prognostic SR-lncRNAs with p values < 0.01. To create a reliable risk predictive model, patients were randomly and equally divided into training ($n = 245$) and testing ($n = 245$) cohorts. The aforementioned 54 SR-lncRNAs were analyzed

using LASSO Cox regression and cross-validated, and 20 prognostic lncRNAs were selected (Fig. 1A, B). Step-by-step multivariate Cox regression analysis further purified 7 lncRNAs for improved clinical utility (Fig. S3). Finally, a prognostic risk model was developed utilizing the correlation coefficient of the expression levels of lncRNAs: risk score = $(0.4295 \times \text{status of OGFRP1}) + (0.3006 \times \text{status of PRKG1-AS1}) + (0.2636 \times \text{status of AL353746.1}) + (0.1115 \times \text{status of SATB2-AS1}) + (-0.0850 \times \text{status of FTO-IT1}) + (-0.3837 \times \text{status of MED4-AS1}) + (-0.4211 \times \text{status of AC090559.1})$.

LUAD patients in the training cohort were categorized as low- or high-risk by median risk score (0.058140853). The survival distribution curve showed a greater death rate for high-risk patients. Heatmap showed that, in the high-risk group, 4 hazardous lncRNAs were highly expressed, while the others were down-regulated (Fig. 1C). The Kaplan–Meier plot demonstrated a lower overall survival rate for the high-risk group (Fig. 1D). Good performance was shown by receiver operating characteristic (ROC) curves with values of 0.77 at 1 year, 0.77 at 3 years, and 0.82 at 5 years for area under the curve (Fig. 1E). Furthermore, compared with previously published prognostic models and clinicopathological risk factors (such as age, gender, and TNM stage), our signature had better prognostic prediction value (Fig. S4). Similar results were found in the testing cohort and the whole cohort (Fig. S5). To make the 7-SR-lncRNAs signature clinically useful, nomograms were created by combining them with or without clinicopathological risk factors (Fig. 1F; Fig. S6A). Calibration curves showed that the 7-SR-lncRNAs-integrated nomogram was more consistent with observation than the simple clinicopathologic nomogram (Fig. S6B, C). The 7-SR-

Peer review under responsibility of Chongqing Medical University.

<https://doi.org/10.1016/j.gendis.2024.101338>

2352-3042/© 2024 The Authors. Publishing services by Elsevier B.V. on behalf of KeAi Communications Co., Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

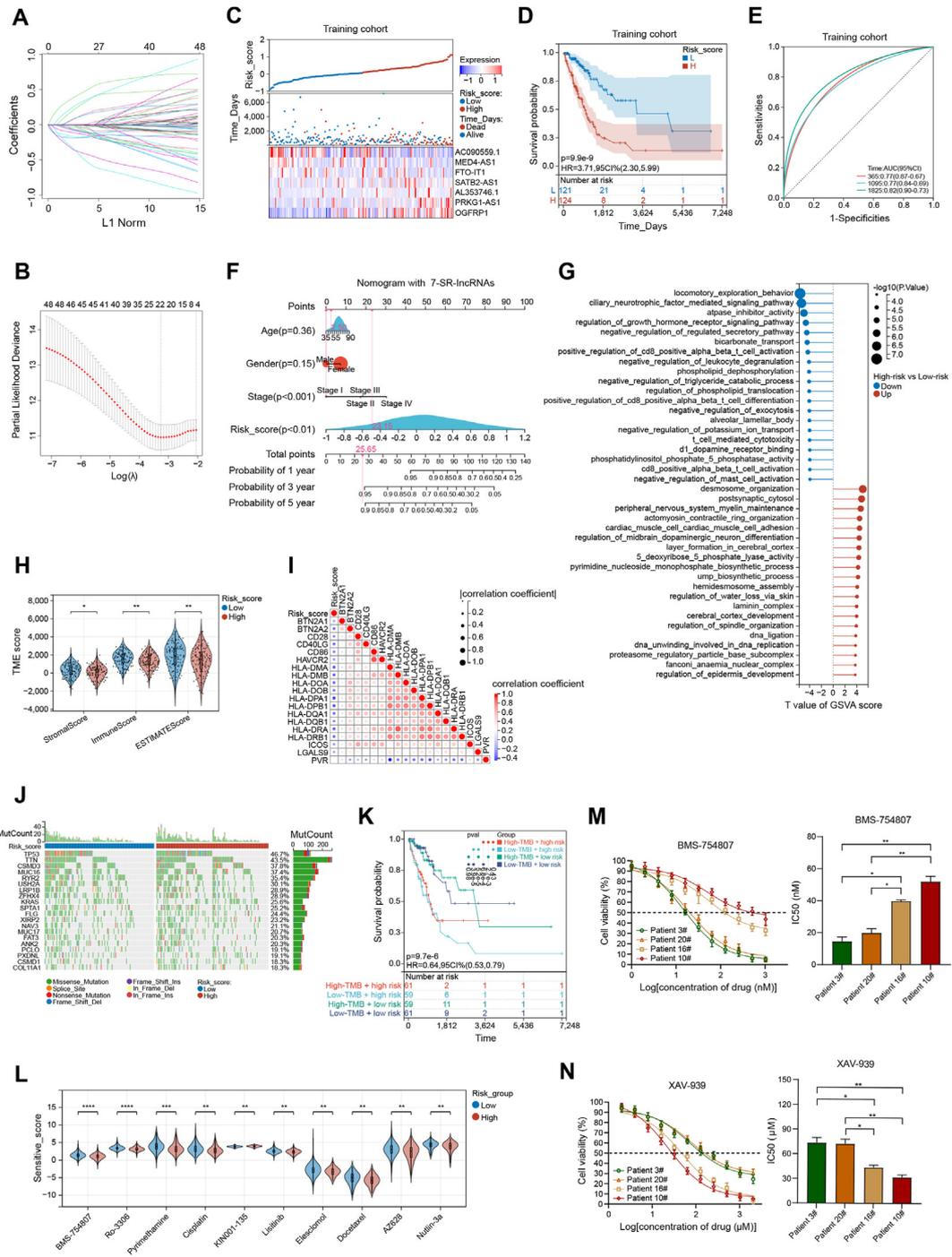


Figure 1 A novel SUMOylation-related lncRNA signature for predicting the prognosis, tumor immune microenvironment, and therapeutic sensitivity of lung adenocarcinoma. **(A)** Profiles of LASSO coefficients for 54 prognostic SR-lncRNAs. **(B)** Ten-fold cross-validation for selecting the LASSO Cox regression model's tuning parameter lambda. The red dots represent the partial likelihood of deviance values, the grey lines represent standard error (SE), and the two vertical dashed lines on the left and right represent optimal values by minimal and 1-SE criteria, respectively. **(C)** The distribution of risk scores, OS status, and expression profiles of the 7 SR-lncRNAs in the training cohort. **(D)** Kaplan–Meier curve for OS in the training cohort. **(E)** The receiver operating characteristic (ROC) curves show the potential of the 7-SR-lncRNAs signature in predicting 1-, 3-, and 5-year OS in the training cohort. **(F)** Nomograms integrated with the 7-SR-lncRNAs signature for predicting OS probability in the training cohort. **(G)** Gene Set Variation Analysis (GSVA) between low- and high-risk groups. **(H)** Comparison of the stromal score, immune score, and ESTIMATE score between low- and high-risk groups. **(I)** The correlation heatmap displays the associations between risk score and 19 differentially expressed immune checkpoint genes. **(J)** The waterfall plot of somatic mutation features in two groups. **(K)** Kaplan–Meier

lncRNAs-integrated nomogram may be superior for LUAD clinical application because the values for its area under the curve at 1, 3, and 5 years were significantly higher (Fig. S6D, E).

Furthermore, numerous approaches were used to investigate the biological activities of the 7-SR-lncRNAs risk model. The Gene Set Variation Analysis (GSVA) found that the two risk groups had differing enrichments of cancer-related pathways like regulation of growth hormone receptor signaling pathway, positive regulation of CD8⁺ alpha beta T-cell activation, and T-cell mediated cytotoxicity (Fig. 1G). In the Gene Set Enrichment Analysis (GSEA) plots, these SR-lncRNAs were enriched in sumoylation, metabolism, and immunology pathways (Fig. S7A–G). Co-expression of protein-coding genes with the 7 SR-lncRNAs was identified using a criterion of $|\text{Pearson correlation coefficient}| > 0.5$ and $p < 0.001$. Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) analyses revealed that those genes were enriched in immune response, Th1/Th2 cell differentiation, and B cell receptor signaling (Fig. S7H, I). In addition to its relevance to classical cancer-related pathways, these findings revealed that the 7-SR-lncRNAs signature was also linked to the immune response.

To study the risk model's role in the tumor immune microenvironment, immune cell infiltration in tumor samples was examined. CIBERSORT, EPIC, MCPCounter, quanTIseq, TIMER, and xCell identified distinct immune cell expression in low- and high-risk groups (Fig. S8A). Using the xCell algorithm, we observed that low-risk samples had more infiltrating immune cells, such as CD4⁺ T-cells, CD8⁺ T-cells, and monocytes (Fig. S8A). Additionally, ESTIMATE showed statistically significant stromal activity variation (Fig. 1H). In light of the development of immune checkpoint inhibitors, we further examined the expression of 79 immune checkpoint genes in low- and high-risk groups,⁴ and found 19 of them were differentially expressed (Fig. S8B). The correlation heatmap showed that these 19 immune checkpoint genes were strongly linked with the risk score (Fig. 1I). These data demonstrated that the 7-SR-lncRNAs signature might assess the tumor immune microenvironment and immune checkpoint gene expression in LUAD patients.

Tumor mutation burden (TMB) can predict immune checkpoint inhibitor responses.⁵ The TMB of low- and high-risk LUAD patients is shown in Fig. 1J. However, the difference in TMB levels between the two groups was not statistically significant ($P = 0.09$; Fig. S9A), and the groups with low- or high-TMB levels had similar survival rates ($P = 0.47$; Fig. S9B). Subgroup analysis stratified samples by TMB status and risk groups was further performed. Intriguingly, some patients with high-risk scores in the low- or high-TMB groups had significantly shorter overall survival

than those with low-risk scores, but no significance was seen between TMB groups with the same risk ratings (Fig. 1K). These findings suggested that the combination of risk group and TMB could provide a more accurate estimate of patient prognosis.

The disparity between the tumor immune microenvironment and TMB may result in differential susceptibility to drug response. We used the “pRRophetic” R package to compare pharmacological efficacy across high- and low-risk groups and found that 45 drugs had statistically significant differences ($P < 0.05$; Fig. S10). Among the top 10 differentially sensitive drugs, BMS-754807, Ro-3306, pyrimethamine, cisplatin, elesclomol, docetaxel, AZ628, and nutlin-3a may be more effective in high-risk patients, and low-risk patients may benefit more from KIN001-135 and lisitinib (Fig. 1L).

A clinical cohort of 80 LUAD patients was created to externally confirm the stability of the 7-SR-lncRNAs signature. Quantitative reverse transcription PCR showed that LUAD tissues expressed more SATB2-AS1, AL353746.1, PRKG1-AS1, and OGFIP1 than neighboring normal tissues, but less AC090559.1, MED4-AS1, and FTO-IT1 (Fig. S11A). In the clinical cohort, the 7 SR-lncRNAs exhibited similar risk scores, overall survival status, and expression profiles to the training cohort (Fig. S11B). The Kaplan–Meier curve and ROC curve indicated that the model could effectively predict prognosis in the clinical cohort as well (Fig. S11C, D). Finally, based on the differing sensitivity profiles (Fig. S10), we validated the sensitivity difference of BMS-754807 and XAV-939 in patient-derived cells obtained from high- and low-risk LUAD clinical samples. The dose–response curves showed that BMS-754807 was more effective in high-risk patients (patients 3# and 20#; Fig. 1M), while low-risk patients (patients 10# and 16#) were more sensitive to XAV-939 (Fig. 1N).

To the best of our knowledge, this is the first study to develop an SR-lncRNAs predictive signature in patients with LUAD and validate it in an independent clinical cohort. This risk model showed good diagnostic accuracy in predicting patient survival outcomes and helped optimize individual treatment strategies. Patient stratification may be improved and novel drugs can be explored in distinct risk groups by making use of differences in features between low- and high-risk groups, such as immune infiltration, TMB, and drug sensitivity. We anticipate eagerly future research into the mechanisms by which these SR-lncRNAs regulate sumoylation and impact patient outcomes.

Ethics declaration

The collection of human tissues was approved by the Medical Ethics Committee of Shanghai Pulmonary Hospital

analysis of OS for patients classified by combing tumor mutation burden (TMB) status and risk score. (L) The top 10 differentially sensitive drugs between low- and high-risk groups. (M) Cytotoxicity assays of BMS-754807 in patient-derived cells (PDCs) constructed from high- (patients 3# and 20#) and low-risk patients (patients 10# and 16#). (N) Cytotoxicity assays of XAV-939 in PDCs constructed from high- (patients 3# and 20#) and low-risk patients (patients 10# and 16#). SR-lncRNAs, sumoylation-related lncRNAs; OS, overall survival.

(Approval number: K23-075Y). Written informed consent was obtained from each patient.

Author contributions

Conception and design: J.C., X.Z., C.C., and D.Z. Collection and assembly of data: J.C., Z.W., and G.L. Data analysis and interpretation: J.C., W.R., B.Y., and H.W. Conducting experiments: J.C., W.R., and Z.W. Manuscript writing and revisions: J.C., X.Z., C.C., and D.Z. Final approval of manuscript: all authors. Accountable for all aspects of work: all authors. All authors read and agreed to the published version of the manuscript.

Conflict of interests

The authors have no conflict of interests to declare.

Funding

This work was supported by the Shanghai Sailing Program (China) (No. 22YF1437600 to J.C.) and the Programs of Shanghai Pulmonary Hospital (China) (No. fkzr2477 to J.C.).

Data availability

All data are available. Please contact us to access it if it is needed.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.gendis.2024.101338>.

References

1. Siegel RL, Miller KD, Wagle NS, Jemal A. Cancer statistics, 2023. *CA Cancer J Clin*. 2023;73(1):17–48.

2. Chaft JE, Rimner A, Weder W, Azzoli CG, Kris MG, Cascone T. Evolution of systemic therapy for stages I-III non-metastatic non-small-cell lung cancer. *Nat Rev Clin Oncol*. 2021;18(9):547–557.
3. Xie M, Yu J, Ge S, Huang J, Fan X. SUMOylation homeostasis in tumorigenesis. *Cancer Lett*. 2020;469:301–309.
4. Hu FF, Liu CJ, Liu LL, Zhang Q, Guo AY. Expression profile of immune checkpoint genes and their roles in predicting immunotherapy response. *Briefings Bioinf*. 2021;22(3):bbaa176.
5. Chan TA, Yarchoan M, Jaffee E, et al. Development of tumor mutation burden as an immunotherapy biomarker: utility for the oncology clinic. *Ann Oncol*. 2019;30(1):44–56.

Jie Cai ^{a,1}, Weizhong Ruan ^{b,1}, Zida Wang ^{c,1}, Gongzhe Liu ^d,
Bei Yang ^a, Hao Wang ^a, Deping Zhao ^{a,*}, Chang Chen ^{a,**},
Xiaogang Zhao ^{a,***}

^a Department of Thoracic Surgery, Shanghai Pulmonary Hospital, School of Medicine, Tongji University, Shanghai 200433, China

^b Department of Thoracic Surgery, Clinical Oncology School of Fujian Medical University, Fujian Cancer Hospital (Fujian Branch of Fudan University Shanghai Cancer Center), Fuzhou, Fujian 350014, China

^c Department of Emergency, Shanghai Pulmonary Hospital, School of Medicine, Tongji University, Shanghai 200433, China

^d Department of Cardiothoracic Surgery, People's Hospital Affiliated to Shandong First Medical University, Jinan, Shandong 271199, China

*Corresponding author.

**Corresponding author.

***Corresponding author.

E-mail addresses: depingzhao_507@sina.com (D. Zhao),
changchen0917@163.com (C. Chen), zxg_shfk@126.com (X. Zhao)

21 October 2023

Available online 28 May 2024

¹ These authors have contributed equally to this work.