

RESEARCH

Open Access



Exploring ribosome biogenesis in lung adenocarcinoma to advance prognostic methods and immunotherapy strategies

Zipei Song^{1†}, Yuheng Wang^{1†}, Miaolin Zhu^{2†}, Pengpeng Zhang³, Zhihua Li¹, Xin Geng¹, Xincen Cao¹, Jianan Zheng^{1*}, Jianwei Tang^{1*} and Liang Chen^{1*} 

Abstract

Background Lung adenocarcinoma (LUAD) presents a considerable danger to human health and has evolved into a major public health concern. Ribosome biogenesis (RiboSis) is a critical process for synthesizing ribosomes, closely associated with cancer initiation, progression, and treatment resistance, potentially serving as a target for future cancer therapies.

Methods Utilizing single-cell RNA sequencing (scRNA-seq) technology, a single-cell atlas of LUAD was delineated, focusing on the analysis of T cell subpopulations. Cells were scored based on the expression patterns of 331 genes associated with RiboSis across different cell types, and monocle2 was employed to analyze the developmental trajectory of CD4⁺ T cells. Employing various machine learning algorithms, a ribosome biogenesis-related signature (RBS) was constructed and compared to 140 published LUAD prognostic models. The relationship between RBS risk scores and various factors in LUAD patients, including prognosis, the tumor immune microenvironment (TIME), responsiveness to immunotherapy, and sensitivity to pharmacological treatments was specifically analyzed. Immunohistochemistry was utilized to validate the expression levels of immune markers in the high- and low- RBS groups, and in vitro experiments were performed to validate the functional role of the pivotal gene KIF23 in the progression of LUAD.

Results Using single-cell analysis, two distinct T cell subtypes were identified: CD8⁺ interferon (IFN) response T cells and CD4⁺ stress response T cells. It was observed that CD4⁺ naive-like T cells exhibit high expression of RiboSis-related genes, with a gradual decrease in RiboSis activity as CD4⁺ T cells develop. Compared to other prognostic models, RBS demonstrated superior performance in prognosis prediction. The low-RBS group exhibited a tumor microenvironment (TME) more favorable for efficient immune monitoring and reaction, higher responsiveness to

[†]Zipei Song, Yuheng Wang and Miaolin Zhu contributed equally to this work and share first authorship.

*Correspondence:

Jianan Zheng

137973781@qq.com

Jianwei Tang

jwtanghaku@126.com

Liang Chen

clbright0909@njmu.edu.cn

Full list of author information is available at the end of the article



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, 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 you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. 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-nc-nd/4.0/>.

immunotherapy, and a better prognosis. Immunohistochemistry confirmed higher expression levels of immune markers in the low-RBS group, while in vitro experiments validated the promoting role of KIF23 in LUAD cell proliferation, migration and invasion.

Conclusion This study delves into the relationship between RiboSis and LUAD cell subpopulations, identifying a potent prognostic biomarker for LUAD. This biomarker aids in assessing immunotherapy efficacy in LUAD patients, ultimately enhancing their prognosis and guiding clinical decision-making.

Keywords Lung adenocarcinoma, Ribosome biogenesis, Single-cell RNA sequencing, Machine learning, Diagnostic biomarkers, Immunotherapy response

Introduction

Lung cancer (LC), characterized by its high mortality rate, is a prevalent malignant tumor [1]. It is broadly classified into two main types: small cell lung cancer (SCLC) and non-small cell lung cancer (NSCLC) [2]. Within the NSCLC subtypes, lung adenocarcinoma (LUAD) and lung squamous cell carcinoma (LUSC) are the most frequent [2]. For most early-stage LUAD patients, early lesion resection through surgery is the most effective treatment [3, 4]. The cure rate for advanced-stage LUAD remains low, with many patients at this stage having lost the opportunity for surgery and can only extend survival through methods such as chemotherapy and targeted therapy [4, 5]. In recent years, immunotherapy, specifically immune checkpoint blockade (ICB) specifically targeting programmed cell death protein 1 (PD-1) and cytotoxic T-lymphocyte associated antigen 4 (CTLA-4) pathways has emerged prominently, which has revolutionized cancer treatment, bringing hope to advanced-stage LUAD patients [6, 7]. However, not all patients respond favorably to immunotherapy [6]. Therefore, identifying specific subsets of LUAD suitable for immunotherapy is a crucial focus in contemporary oncology [8].

Ribosome biogenesis (RiboSis) is a complex and crucial cellular process that initiates in the nucleolus and concludes in the cytoplasm, generating ribosomes essential for cell growth and proliferation [9, 10]. Ribosomes are the mRNA translation machinery responsible for protein synthesis [11]. RiboSis involves multiple steps, including rRNA transcription, rRNA splicing, rRNA modifications, ribosome assembly, and output of ribosomal precursors [9, 10, 12]. Recent studies indicate that ribosomes in tumors exhibit heterogeneity, giving rise to the concept of “cancer ribosomes” [9]. Increasing evidence suggests that RiboSis is closely associated with cancer initiation, metastasis, and therapy resistance, thereby positioning it as a potential target for the development of innovative cancer treatment approaches [9, 12–14]. Existing research has highlighted the crucial role of RiboSis in the progression of NSCLC [15, 16]. Wang et al. discovered that SOD1 in the cell nucleus promotes the proliferation of KRAS-driven NSCLC cells by regulating ribosome

biogenesis and proliferation [15]. Lian et al. found that T cell differentiation protein 2 enhances mTOR-mediated ribosome biogenesis to stimulate NSCLC cell proliferation [16]. However, the impact of RiboSis on LUAD metastasis, treatment resistance, and clinical outcomes has yet to be fully elucidated.

In recent years, the continuous advancement of single-cell RNA sequencing (scRNA-seq) technology and bioinformatics has pioneered a new field in the study of LUAD [17]. ScRNA-seq aids in revealing cellular heterogeneity, enabling precise identification and differentiation of diverse cell subpopulations, and discovering specific subgroups that have a vital impact on the emergence and evolution of LUAD [18]. Identifying biomarkers for LUAD helps in evaluating patients' disease status, predicting disease progression, and assessing treatment responses [19, 20]. Biomarkers play a significant role in personalized medicine, facilitating treatment planning, optimizing therapeutic outcomes, and ultimately augmenting both survival rates and the quality of life for patients with LUAD [19].

In this study, we utilized a scRNA-seq dataset to identify cell subpopulations in LUAD samples and extensively investigated the roles of RiboSis in various cell types. Subsequently, we developed a ribosome biogenesis-related signature (RBS) aimed at improving the prognosis of LUAD patients, assessing immunotherapy efficacy, and providing new insights for the purpose of clinical diagnosis and management of LUAD.

Materials and methods

Data collection and initial processing

The data derived from scRNA-seq for nine primary LUAD samples, were obtained from the Gene Expression Omnibus (GEO) database with the accession number GSE189357, accessible at <https://www.ncbi.nlm.nih.gov/geo/> [21]. These data were retrieved specifically for subsequent single-cell analysis. Additionally, bulk RNA sequencing (bulk RNA-seq) data, along with its corresponding clinical information of 1,498 LUAD patients were obtained from various public databases for the purpose of constructing and validating RBS. The sources of these data include The Cancer Genome Atlas

(TCGA) database, accessible via its website at <https://portal.gdc.cancer.gov/>, provides a wealth of data, including bulk RNA-seq data, copy number variation (CNV) data, mutation data, and clinical data of 596 LUAD patients. Furthermore, GEO databases provided additional data, with accession numbers GSE30219 ($n=83$) [22]; GSE31210 ($n=246$) [23]; GSE42127 ($n=131$) [24]; and GSE68465 ($n=442$) [25], encompassing RNA-seq data and clinical data of 902 LUAD patients. To facilitate comparative analysis, bulk RNA-seq data of 288 normal lung tissue samples were downloaded from the Genotype-Tissue Expression (GTEx) database (website: <https://gtexportal.org/home/>). Two immunotherapy cohorts specific to LUAD, namely the OAK dataset ($n=241$) [26] and the POPLAR dataset ($n=55$) [27], were utilized in this study to assess the predictive capability of the model in forecasting the prognosis of patients undergoing immunotherapy. Additionally, three pan-cancer immunotherapy cohorts, which included the Melanoma-GSE91061 dataset ($n=109$) [28], Melanoma-phs000452 dataset ($n=153$) [29], and Melanoma-PRJEB23709 dataset ($n=91$) [30], were also employed for this purpose. A cohort encompassing 21 distinct types of tumors, constituting a pan-cancer dataset, was utilized to evaluate the broad applicability of the model (accessible at <https://xena.ucsc.edu/>). In the subsequent analyses conducted, samples that lacked clinical information and with a total survival period (OS) of 0 were excluded. To ensure data comparability, all bulk RNA-seq data were normalized to transcripts per million (TPM) except for differential analysis, which used counts data format. To address potential batch effects, the “combat” function from the “sva” R package was utilized for potential batch effect removal [31]. Prior to analysis, a log2 transformation was applied to all bulk RNA-seq data to ensure that the data were on a more appropriate scale for statistical analysis.

To elucidate the role of RiboSis in the biology and prognosis prediction of LUAD, we constructed a gene set related to RiboSis consisting of 331 genes based on the Gene Ontology (GO) term of the Molecular Signatures Database (MSigDB) and the characterization by Nerurkar et al. (Table S1) [32, 33].

Preliminary analysis of scRNA-seq data

In the comprehensive analysis of the GSE189357 dataset, the scRNA-seq data from nine LUAD samples were processed using the Seurat package, version 5.0.1, within the R software environment, version 4.3.0 [34]. Initially, a rigorous cell quality control procedure was implemented, which involved the systematic retaining of genes and cells of high quality based on a variety of criteria. These criteria encompassed genes that were expressed in a minimum of three cells, the number of genes expressed per cell ranging between 300 and 7000, the expression levels

of mitochondrial genes kept below 10%, and the expression levels of red blood cell-related genes maintained under 3%, and cells with less than 100,000 unique molecular identifiers (UMIs). After following the quality control procedures, 114,341 high-quality cells were chosen for a more in-depth examination. In the next phase, a series of data preprocessing steps were executed, including normalization, scaling, and the identification of highly variable genes. These tasks were efficiently handled using the specific functions named “NormalizeData”, “ScaleData”, and “FindVariableFeatures”. To further enhance the quality of the data, potential batch effects were addressed by utilizing the “RunHarmony” function, the “RunPCA” function was applied to identify anchors, subsequently employing the “RunUMAP” function for dimensionality reduction [35]. Clustering analysis was executed with the “FindNeighbors” and “FindClusters” functions (resolution = 0.5). Marker genes were pinpointed by contrasting cells within specific clusters against all other clusters using the “FindAllMarkers” function. Annotation of data was done according to representative marker genes for the primary cell types found in LUAD. RiboSis activity scores for individual cells were computed using the “AUCell” and “AddModuleScore” R packages [36, 37]. Cells were then divided into two groups: those with high RiboSis-AUC scores and those with low RiboSis-AUC scores, using the median AUC value as the cutoff. The “cellchat” R package was employed to assess the variations in cell communication between the high and low RiboSis-AUC groups [38].

Construction of T cell atlas and trajectory analysis

From the annotated data, we selected 39,149 T cells and 6,071 NK cells for further investigation. The data processing methods were similar to the previous steps, involving normalization, identification of highly variable genes, batch effect removal, dimensionality reduction, and clustering. When performing clustering using the “FindClusters” function here, the resolution was set to 1.0. Subsequently, cell clusters were annotated conducted on the basis of differentially expressed genes (DEGs) and typical immune markers. RiboSis activity scores were computed for each cell using the “AddModuleScore” R package. Cell communication within T cell subpopulations was assessed using the “cellchat” R package. Furthermore, Monocle2 (version 2.30.0) was employed to infer the cellular lineage trajectory of CD4⁺ T cells [39, 40]. After dimensionality reduction and cell ordering, differentiation trajectories were deduced using Monocle’s predefined settings. Signature genes with q -values < 0.001, calculated using the “differentialGeneTest” function, were used for RiboSis activity scoring of the inferred temporal results. Finally, enrichment analysis

of the inferred pseudotime results was conducted using “clusterProfiler” and the “org.Hs.eg.db” R package.

Screening key RiboSis-related genes

Utilizing the “FindMarkers” function, we were able to pinpoint DEGs between the high and low RiboSis-AUC groups, employing the criteria of $|\log_2$ fold change (FC)| > 0.7 and $p < 0.05$. Following this, a Spearman correlation analysis was performed, ultimately selecting the top 150 genes that demonstrated the strongest correlation with RiboSis. By merging these highly correlated genes with DEGs, we obtained a comprehensive list of 830 RiboSis-related genes for further analysis. The bulk RNA-seq data originated from 288 normal lung tissues, sourced from the GTEx database and merged with the TCGA dataset. Normalization was performed applying a $\log_2(\text{count} + 1)$ transformation and mitigated batch effects. With the normalized and cleaned dataset, we focused our attention on the 830 RiboSis-related genes, resulting in 369 DEGs between LUAD and normal tissues ($P < 0.05$, $|\log_2\text{FC}| > 0.8$). Subsequently, univariate Cox regression analysis was carried out using the TCGA-LUAD dataset, leading to the identification of 83 RiboSis-related genes that demonstrated prognostic significance ($P < 0.05$).

Construction of RBS through combinations of machine learning algorithms

To establish a comprehensive and reliable prognostic signature known as RBS, we integrated 10 machine learning (ML) algorithms: Lasso, elastic net (Enet), stepwise Cox, partial least squares regression for Cox (plsRcox), supervised principal components (SuperPC), random survival forest (RSF), CoxBoost, generalized boosted regression modeling (GBM), Ridge, and survival support vector machine (survival-SVM) [41–43]. The TCGA-LUAD dataset served as the training set, whereas GSE30219, GSE31210, GSE42127, and GSE68465 with batch effects mitigated, were chosen as validation sets. The outline of RBS construction is as follows: (a) Leveraging a machine learning framework, 81 prognostic models were constructed using 101 algorithm combinations based on key genes. (b) The performance of these models was assessed using four validation datasets. (c) The concordance index (C-index) was computed for each model across all four validation datasets, and the one with the highest average C-index was deemed the optimal model. (d) Further analysis was conducted using the selected optimal model to pinpoint the involved genes, followed by risk score computation for each patient. Subsequently, within each dataset, patients were categorized into high-RBS group and low-RBS group based on the median risk score. This rigorous process guarantees the creation of a validated, reliable, and robust signature across multiple datasets.

Assessment of RBS and construction of a nomogram

Kaplan-Meier survival curves were produced using the “survminer” R package to investigate the disparities in overall survival (OS) and progression-free survival (PFS) between the high and low RBS groups. For evaluating the precision of RBS in predicting LUAD patients’ prognosis, receiver operating characteristic (ROC) analysis was executed with the “timeROC” R package, calculating the area under the curve (AUC). Dimensionality reduction of data was executed via principal component analysis (PCA), visualizing sample distribution differences between groups. A clinical correlation circle plot was generated to explore the relationship between RBS and clinical characteristics. The “CompareC” R package facilitated calculating and comparing the C-index of RBS and clinical features. Following a PubMed literature search, we obtained 140 signatures designed to forecast the prognosis of patients with LUAD. Subsequently, we compared the C-index of the RBS with the C-index of these signatures across multiple datasets. Finally, by combining RBS with TNM staging, a nomogram was constructed to forecast the survival rates for patients with LUAD. To comprehensively evaluate the predictive accuracy and overall performance, we utilized calibration curves, decision curve analysis (DCA), and ROC curves as graphical representation tools.

Mutation landscape

Using GISTIC 2.0 analysis, genomic alterations including recurrent amplifications and deletions were identified. Additionally, the “maftools” R package was utilized to calculate the tumor mutational burden (TMB) [44]. LUAD patients were divided into two groups based on their TMB, using the median TMB score as the cutoff point. Finally, integrating clinical data from TCGA-LUAD dataset, a survival analysis was performed to assess and compare the prognostic outcomes between these two patient groups.

Pathway enrichment

Utilizing the MsigDB database and R software packages such as “GSVA”, “ClusterProfiler”, “org.Hs.eg.db” and “GseaVis”, gene set variation analysis (GSVA) and gene set enrichment analysis (GSEA) were conducted [45]. Enrichment analysis of DEGs between the high and low RBS groups was carried out utilizing the comprehensive tools available on the Metascape website. (<https://metascape.org/>) [46]. Enrichment scores for each sample were calculated based on enriched pathways and were then visualized through the application of t-distributed stochastic neighbor embedding (tSNE) dimensionality reduction.

Immune infiltration

The TIMER2.0 database (accessible at <http://timer.comp-genomics.org/timer/>) consolidates findings from multiple

algorithms and provides a comprehensive summary of immune cell infiltration levels in TCGA datasets [47]. We leveraged this resource to examine variations in immune cell infiltration between the high and low RBS groups. Utilizing the “ESTIMATE” R package, we computed the immune score and stromal score, and ESTIMATE score for each LUAD patient in the TCGA-LUAD dataset, and conducted correlation analysis to assess their association with RBS risk scores. To evaluate potential differences in the Tumor Immune Microenvironment (TIME) between high- and low- RBS groups, a total of 8 unique algorithms were employed, including TIMER, xCell, CIBERSORT, CIBERSORT-ABS, quanTIseq, MCP-counter, EPIC, and ssGSEA [47–49], the correlation between the RBS risk score and the cancer-immunity cycle was subsequently investigated. Ultimately, a comparative assessment of the expression of immune checkpoint (IC) molecules and immune modulators was conducted.

Immunotherapeutic response prediction

The relationship between the tumor immune evasion score and the RBS risk score was evaluated utilizing the Tumor Immune Dysfunction and Exclusion (TIDE) website, accessible at <http://tide.dfci.harvard.edu/>. Additionally, the Immune Phenotype Score (IPS) for patients with LUAD was retrieved from The Cancer Immunome Atlas (TCIA) database (<http://tcia.at/home>) to evaluate their immunogenicity and predict their response to ICB therapy drugs. We employed the model's formula to calculate the RBS scores for five immunotherapy cohorts (two LUAD immunotherapy cohorts (OAK, POPLAR) and three pan-cancer cohorts (Melanoma-GSE91061, Melanoma-phs000452, and Melanoma-PRJEB23709)) and conducted survival analysis.

Predicting drug sensitivity and screening small molecule drugs

The data for drug sensitivity were acquired from various reputable sources: the Cancer Therapeutics Response Portal database (CTRP 2.0, accessible via the URL: <https://portals.broadinstitute.org/ctrp.v2.1/>), the PRISM database (accessible via the URL: <https://www.theprismlab.org/>), the Cancer Drug Sensitivity Genomics of Cancer Cell Lines (GDSC) database (accessible via the URL: <https://www.ancerrxgene.org/>). In parallel, the cell line expression profile data were extracted from the Cancer Cell Line Encyclopedia (CCLE) database (accessible via the URL: <https://sites.broadinstitute.org/ccle/>). A ridge regression model was constructed leveraging the transcriptome data provided by TCGA. By performing a comprehensive analysis, the Spearman correlation between the RBS risk score and the AUC score was calculated (CTRP 2.0: Spearman's $r < -0.4$; PRISM: Spearman's $r < -0.55$), a total of 24 small molecular drugs were identified. To further

understand the sensitivity of high and low RBS groups to common anticancer drugs, Utilizing the GDSC database, drug sensitivity data was obtained and the “pRRophetic” R package was employed to compute the half-maximal inhibitory concentration (IC50) values.

Constructing molecular docking models

Obtain the two-dimensional (2D) structures of SB-743,921 and Ispinesib from the PubChem database (accessible via the URL: <http://pubchem.ncbi.nlm.nih.gov/>). After successfully acquiring these 2D structures, proceed to input them into the ChemOffice 20.0 software program in order to generate their corresponding three-dimensional (3D) structures. Following this step, the next task involves retrieving the protein target sequences of KIF23 (Uniprotkb: Q02241) and TPX2 (Uniprotkb: Q9ULW0) from the Uniprot database (accessible via the URL: <https://www.uniprot.org/>), and construct protein crystal models by utilizing the AlphaFold tool, which is available at the web address <https://alphafold.ebi.ac.uk>. After constructing the protein crystal models, perform energy minimization on the compounds using the Molecular Operating Environment 2019 (MOE 2019) software. Subsequently, preprocess the target proteins to identify active pockets. Conclude this step by performing molecular docking using MOE 2019, with a specified number of iterations set at 50. Evaluating the binding activities of the compounds based on the binding energies and visualize the results using PyMOL 2.6.0 and Discovery Studio 2019 software.

Clinical specimen collection and RNA sequencing

The process of collecting tissue samples has been granted ethical approval from the Medical Ethics Committee of the First Affiliated Hospital of Nanjing Medical University. On the surgery day, 10 LUAD samples, which have been categorized into three distinct groups: AIS, MIA, or IAC by pathology experts, are procured and dispatched to Oncocare Inc. located in Suzhou, China, for the purpose of undergoing RNA sequencing. We collected transcriptome sequencing data from 10 LUAD samples and calculated their respective RBS risk scores. Based on these scores, we classified a sample with the highest RBS scores into the high-RBS group, and a sample with the lowest scores into the low-RBS group.

Immunohistochemistry (IHC)

Paraffin-embedded tissue sections at 37°C underwent an incubation process for a duration of 120 min. The primary antibodies utilized were anti-CD4 (Catalog number: UMAB64; ZSGB-BIO, Beijing, China), anti-CD8 (Catalog number: SP16; ZSGB-BIO, Beijing, China), and PD-L1 (Catalog number: MXR003; MXB, Fuzhou, China). Subsequent to this primary step, horseradish

peroxidase-conjugated secondary antibodies were introduced and allowed to incubate for an additional period of 30 min. The temperature conditions were maintained throughout. Following this, the sections underwent staining with DAB (3,3'-diaminobenzidine) and were subsequently counterstained with haematoxylin to enhance visualization.

Cell lines culture

We acquired the 16HBE human bronchial epithelial cell line, the A549 and the H1299 human LUAD cell line from the Institute of Biochemistry and Cell Biology, Chinese Academy of Sciences. It was necessary to culture all the cells in DMEM medium (Gibco; Thermo Fisher Scientific, Inc.) containing 10% fetal bovine serum (FBS) and 1% penicillin-streptomycin (Gibco; Thermo Fisher Scientific, Inc.). The culture conditions were maintained to ensure optimal cell growth and viability. This involved setting the incubator temperature at 37 °C, maintaining a CO₂ concentration of 5% and ensuring a humidity level of 95%.

RNA extraction and qRT-PCR

Total RNA was extracted from the cells utilizing the TRIzol® reagent, which was sourced from Invitrogen; a brand belonging to Thermo Fisher Scientific, Inc. according to the manufacturer's protocol. Reverse transcription was subsequently carried out with the PrimeScript Reverse Transcription Kit (Takara Bio, Inc.). Quantitative real-time polymerase chain reaction assays (qRT-PCR) was performed using the Archimed-X4 Real-Time PCR System (Applied ARCHIMED X; RocGene Technology, Inc.). The specific PCR conditions employed during this process comprised an initial step involving denaturation, which was conducted at a temperature heated to 95 °C for 30 s during the initial denaturation phase, succeeded by 40 cycles of denaturation at 95 °C for 10 s and annealing at 60 °C for 30 s. Relative mRNA expression levels were quantified employing the 2- $\Delta\Delta C_t$ method, utilizing GAPDH as the reference gene for normalization. Triplicates were conducted for each reaction. Below are the primers utilized for the genes investigated in this study:

GAPDH.

Sequence (5' -> 3').

Forward Primer TACCCATTTGAATCGTGAGTCA.

Reverse Primer CTCTGGTCCGGTTAGTTCTTTC.

KIF23.

Sequence (5' -> 3').

Forward Primer AGTCAGCGAGAGCTAAGACAC.

Reverse Primer GGTTGAGTCTGTAGCCCTCAG.

Transfection of plasmid DNA and small interfering RNA

The KIF23 cDNA was inserted into the pcDNA expression vector. Transfection of plasmids was achieved employing the X-tremeGEN™ HP reagent for DNA transfection sourced from Nanjing, China, whereas small interfering RNA (siRNA) transfection was facilitated by Lipo2000 reagent from Invitrogen, USA, both conducted in accordance with the manufacturers' protocols. Coverslips positioned in six-well plates were inoculated with A549 cells, and transfection with either plasmid or siRNA was carried out 48 h post-seeding. The specific target sequences utilized are detailed as follows:

si-KIF23-1# Sequence (5' -> 3').

Sense GACUAUAUCUAGAUAUGUCU.

Antisense ACAUGAUCUAGAUAUAGUCUU.

si-KIF23-2# Sequence (5' -> 3').

Sense GAAGUGAUCAAUAAUACAACU.

Antisense UUGUAUUAUUGAUCACUUCUA.

Cell counting Kit-8 (CCK-8) assay

We performed the Cell Counting Kit- 8 (CCK-8) assay utilizing CCK-8 solution sourced from Dojindo Laboratories, Inc., adhering strictly to the manufacturer's guidelines. Cells were plated in 96-well plates at a density of 2×10^3 cells per 100 μ l. Subsequently, CCK-8 solution was introduced into the wells at various time points: 0, 24, 48, 72, 96, and 120 h. Cell viability was evaluated after an additional 2-hour incubation period, with absorbance utilized a microplate reader to measure at a wavelength of 450 nm.

Colony formation assay

In the colony formation assay, LUAD cells were plated at a density of 1×10^3 cells per well within six-well plates. These plates were then incubated at a temperature of 37°C with an atmosphere containing 5% carbon dioxide, for a period of 14 days. During this incubation period, the growth medium was replaced every four days to ensure optimal conditions for colony formation. Once the incubation period was completed, the colonies were fixed using a 4% solution of paraformaldehyde. Subsequently, they were stained with a 0.1% solution of crystal violet, which was sourced from the reputable Beyotime Institute of Biotechnology. Colonies containing more than 50 cells were meticulously counted in each well to ensure accuracy. The experiments were conducted in triplicate.

Wound-healing assay

Once the LUAD cells reached a 90% confluence, they were plated into 6-well dishes. A linear scratch wound

was then created across the monolayer using the tip of a 20 μ l pipette. After culturing the cells in a medium devoid of FBS, the monolayer was gently rinsed with PBS to eliminate debris and any floating cells. Images of each wound were documented using an inverted microscope immediately after wounding and again after 24 h.

Transwell assay

For the LUAD cell invasion and migration assay, a Transwell chamber (8 μ m pore size; Corning, Inc.) featuring 24 pores with a diameter of 8 mm was employed. The matrix gel invasion analysis necessitates the use of Transwell membranes (Sigma Aldrich) that are pre-coated with matrix gel. In each upper Transwell chamber, a total of 4×10^4 cells were carefully cultured and maintained in precisely 300 μ l of serum-free medium, ensuring optimal conditions for cellular growth and viability. Simultaneously, the lower chamber was filled with 700 μ l of Dulbecco's Modified Eagle Medium (DMEM), which was further supplemented with 10% fetal bovine serum (FBS). Following a 16-hour incubation period, cells on the upper surface were removed and stained with methanol and 0.1% crystal violet after migrating through the membrane to reach the lower surface. Photographic records were captured utilizing an advanced inverted microscope, specifically manufactured by the reputable company Olympus, in Tokyo, Japan.

Statistical analysis

The initial data processing, coupled with statistical analysis and visualization of results, was carried out utilizing R software (version 4.3.0) and Perl software (version 5.30.0). Both Pearson and Spearman correlation analyses were utilized to evaluate the relationship between two continuous variables. For categorical variables, comparisons were made using the Chi-square test. When dealing with continuous variables, either the Wilcoxon rank-sum test or the t-test was utilized for comparisons. All statistical tests were conducted with a two-faceted approach, with statistical significance ascertained by a P-value < 0.05.

Results

The characterization of RiboSis-related genes in LUAD scRNA-seq data

Figure 1 depicts the workflow of this study. Initially, scRNA-seq data from the GSE189357 dataset underwent cellular quality control, resulting in the selection of 114,341 high-quality cells for subsequent analysis. After mitigating batch effects across samples, the cellular distributions of the 9 samples became relatively consistent (Fig. S1A). Following that, dimensionality reduction was carried out using the "RunUMAP" function, clustering all cells into 24 cell clusters (Fig. 2A). Annotation based on typical marker genes of major LUAD cell types was

carried out, resulting in the identification of 9 cell subtypes comprising epithelial cells, endothelial cells, T lymphocytes, NK cells, B lymphocytes, plasma cells, myeloid cells, mast cells, and fibroblasts (Fig. 2B). A bubble plot displayed the typical marker genes of major LUAD cell types (Fig. 2C), a UMAP plot depicted the expression distribution of these marker genes in the data (Fig. 2D), and a proportional plot illustrated the differences in cellular type distributions among the 9 samples (Fig. S1B).

Subsequently, based on the expression levels of 331 RiboSis-related genes in the dataset, the RiboSis activity scores for each cell were calculated. The results indicated that B lymphocytes and T lymphocytes exhibited higher RiboSis activity, while endothelial cells, plasma cells, and epithelial cells showed lower RiboSis activity (Fig. 2E, Fig. S1C). Cells were classified into high RiboSis-AUC and low RiboSis-AUC groups based on their average AUC scores (Fig. 2F). Cell communication variances between the two groups were then assessed, the results revealed that the high RiboSis-AUC group exhibited increased quantity and intensity of cell communication (Fig. 2G, Fig. S1D). Additionally, we observed that the high RiboSis-AUC group demonstrated higher incoming and outgoing interaction strengths compared to the low RiboSis-AUC group (Fig. 2H). Furthermore, most signaling pathways showed elevated expression levels in the high RiboSis-AUC group (Fig. 2I, J). Bubble plots illustrated that the high RiboSis-AUC group displayed stronger intercellular communication through a greater number of ligand-receptor pairs (Fig. 2K). Specifically, interactions between myeloid cells and epithelial cells, as well as endothelial cells, were upregulated for ligand-receptor pairs such as RETN-CAP1 and LGALS9-CD44. In summary, these findings suggest a positive correlation between RiboSis activity in cells and the intensity of intercellular communication.

Construction of T cell atlas and evaluation of ribosis activity

Tumor-infiltrating T cells (TILs) are a critical component of TIME and represent a major subset responsive to immunotherapy in LUAD, exhibiting higher RiboSis activity (Fig. 2E). Therefore, we conducted in-depth analysis of T cells and constructed a T cell atlas for LUAD. Following the removal of batch effects between samples, the distribution of T and NK cells across the 9 samples was relatively consistent (Fig. S1E). Through dimensionality reduction and clustering, 20 cell clusters were identified (Fig. 3A). Subsequently, based on DEGs and typical immune markers, we defined 11 T cell subtypes and 1 NK cell subtype (Fig. 3B) [50].

Firstly, we distinguished NK cells (C4, C11, and C19) characterized by high expression of genes such as NKG7, GNLY, and KLRD1. Next, we identified $\gamma\delta$ T cells (Tgd;

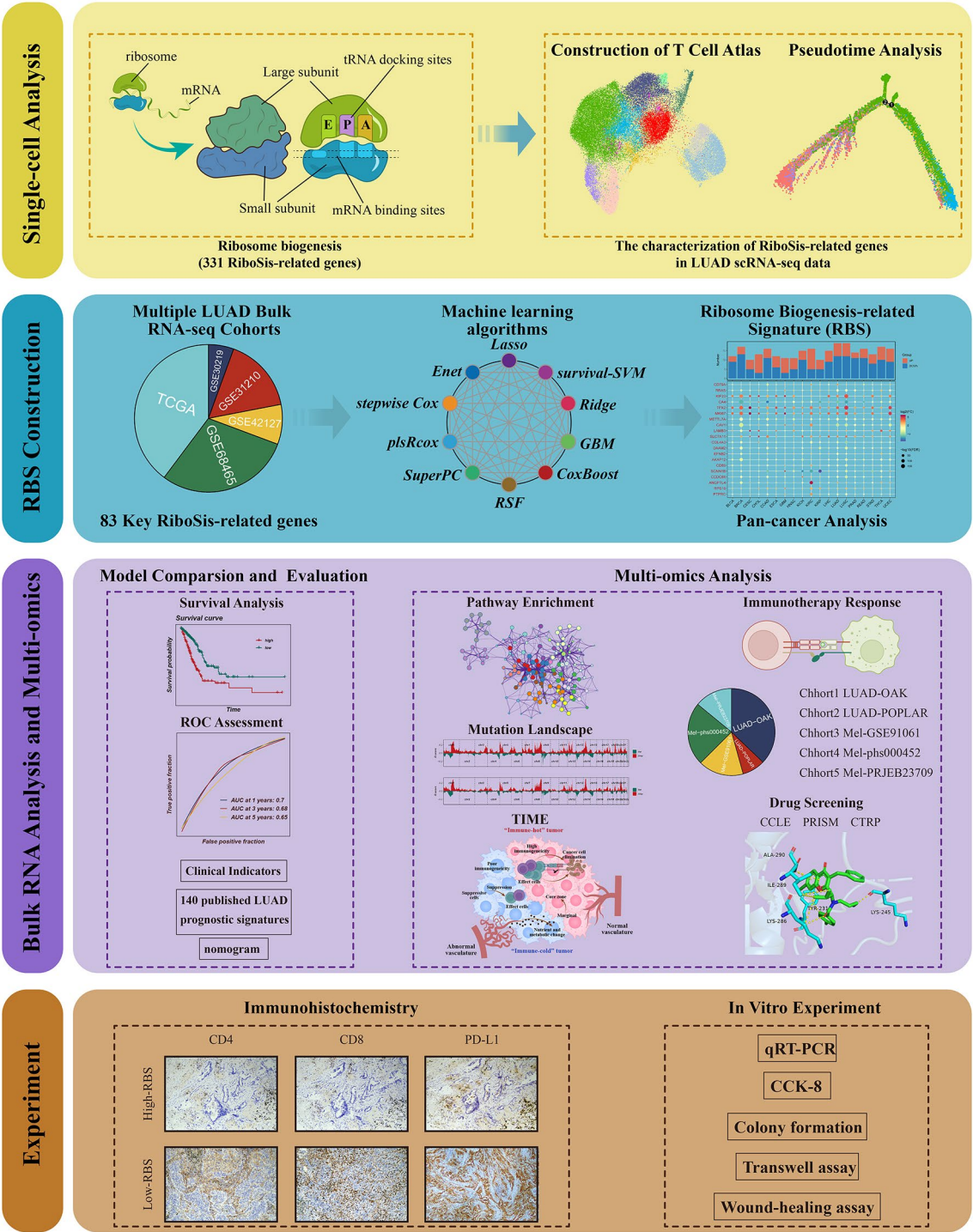


Fig. 1 The flowchart of this study. This figure illustrates the workflow of the study exploring the role of RiboSis in LUAD. The study utilized machine learning to construct RBS for predicting prognosis and immunotherapy response in LUAD patients. Immunohistochemistry experiments confirmed elevated expression levels of immune markers in the low RBS group, and in vitro experiments validated the promoting role of KIF23 in LUAD cell proliferation, migration, and invasion. RiboSis, ribosome biogenesis; LUAD, lung adenocarcinoma; RBS, ribosome biogenesis-related signature; ROC, receiver operating characteristic; TIME, tumor immune microenvironment; qRT-PCR, quantitative real-time polymerase chain reaction assays; CCK-8, cell counting kit-8 experiment

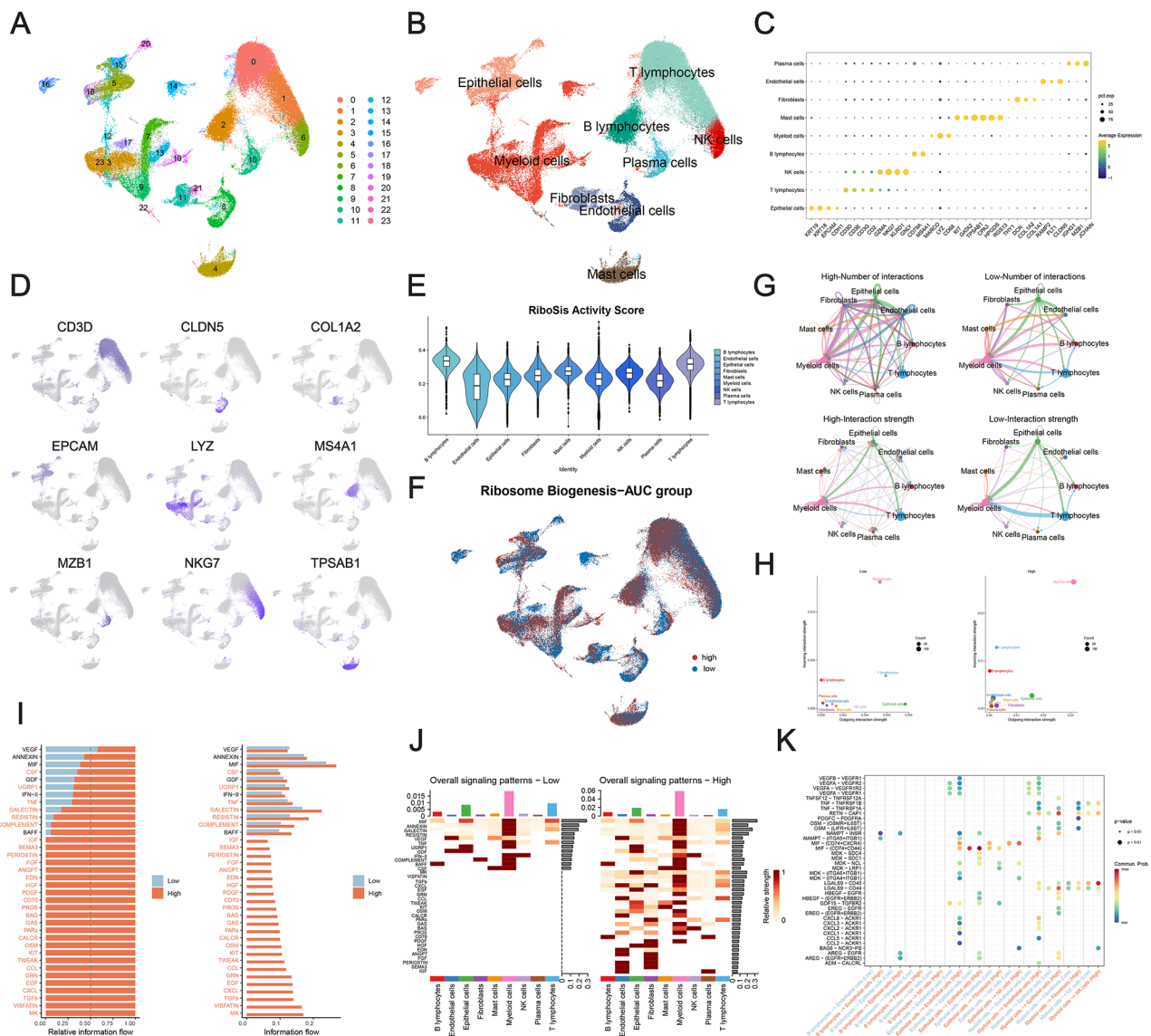


Fig. 2 Characterization of RiboSis-related genes in LUAD scRNA-seq data. **(A)** Following dimensionality reduction and clustering, cells from 9 LUAD samples were grouped into 24 cell clusters. **(B)** Cell annotation yielded 9 cell subtypes. **(C)** Bubble plot displaying typical marker genes for the 9 cell subtypes. **(D)** Expression distribution of 9 typical marker genes across samples. **(E)** Violin plots comparing RiboSis activity scores of the 9 cell subtypes. B cells and T cells exhibit higher RiboSis activity scores, while endothelial cells, plasma cells, and epithelial cells show lower RiboSis activity scores. **(F)** UMAP plot illustrating the cell distribution of high and low RiboSis-AUC groups. **(G)** Network plot comparing cell communication quantity and intensity between high and low RiboSis-AUC groups. **(H)** Contrasting incoming and outgoing interaction strengths among different cell subtypes between high and low RiboSis-AUC groups. **(I, J)** Comparison of signaling pathway strengths between high and low RiboSis-AUC groups. **(K)** Upregulated and downregulated ligand-receptor pairs in different cell subtypes of high and low RiboSis-AUC groups

C13) showing high expression of TRDC and cytotoxicity-related genes. Subsequently, we differentiated CD8⁺ T cells (high expression of CD8A, CD8B, etc.) and CD4⁺ T cells (high expression of CD4, CD40LG, etc.). For CD8⁺ T cells, based on the characteristic high expression of cytotoxicity-related genes (GZMA, GZMB, GZMH, GZMK, etc.), we defined CD8⁺ effector T cells (CD8_TEFF; CD8_TEFF_GZMK: C1 and CD8_TEFF_GZMH: C8). Furthermore, according to the feature of high expression of exhaustion-related genes (FASLG, PDCD1, LAG3, etc.),

we defined CD8⁺ exhausted T cells (CD8_TEX; C16). For CD4⁺ T cells, based on the characteristic high expression of naive genes (CCR7, SELL, TCF7, etc.), we defined CD4⁺ naive-like T cells (CD4_TN; C0, C2, C6, C7, and C17). We also defined CD4⁺ helper T cells (CD4_TFH; C12), characterized by high expression of CXCL13, TOX, ICOS, etc. CD4⁺ regulatory T cells (CD4_Treg; C9 and C10), demonstrating high expression of FOXP3, IL2RA, TNFRSF4, and others. For memory T cells, we defined CD8⁺ resident memory T cells (CD8_TRM; C3 and C15;

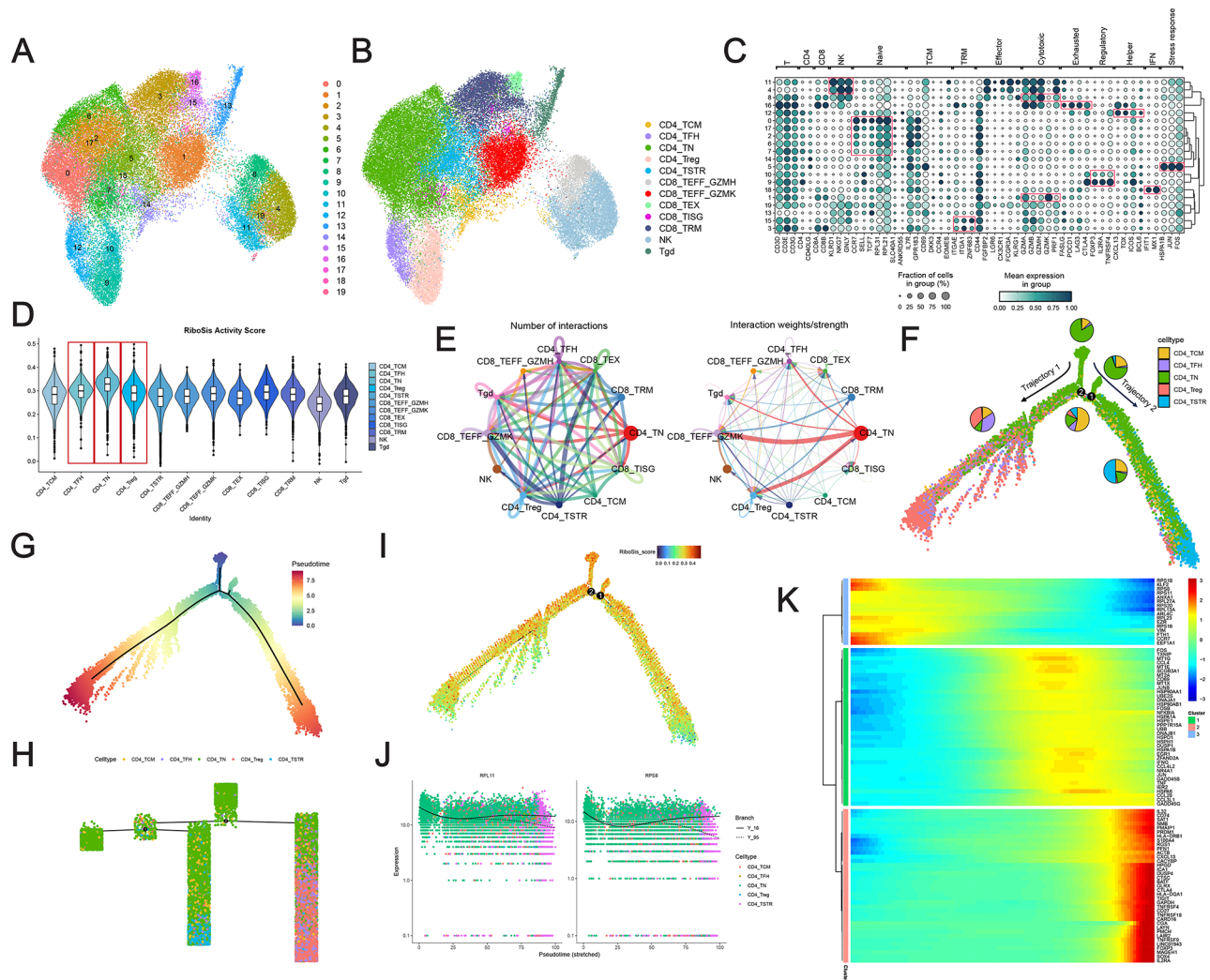


Fig. 3 Construction of T cell atlas and evaluation of RiboSis activity. **(A)** Dimensionality reduction clustering was performed on T cells and NK cells, identifying 20 cell clusters. **(B)** Cell annotation defined 11 T cell subtypes and 1 NK cell subtype. **(C)** Bubble plot illustrating the association between typical marker genes of various T cell subtypes and the 20 cell clusters. **(D)** Violin plots comparing RiboSis activity scores of 12 T/NK cell subtypes. Higher RiboSis activity was observed in CD4_TN, CD4_TFH, CD8_TISG, and CD4_Treg. **(E)** Cell communication quantity and intensity among the 12 T/NK cell subtypes. **(F-H)** Developmental trajectories of CD4+ T cells inferred by Monocle2. **(I)** RiboSis activity scoring based on pseudo-time results. **(J)** Expression changes of two hallmark RiboSis-related genes, RPL11 and RPS8, as differentiation progresses. **(K)** Heatmap displaying expression patterns of genes with the highest differential expression at various stages of CD4+ T cell differentiation

high expression of ITGAE, ITGA1, ZNF683, etc.) and CD4⁺ central memory T cells (CD4_TCM; C14). Additionally, we identified two distinct T cell subtypes. CD8⁺ interferon (IFN) response T cells (CD8_TISG; C18) exhibited high expression of interferon-stimulated genes (IFIT1, MX1, etc.). CD4⁺ stress response T cells (CD4_TSTR; C5) were characterized by the abundant expression of heat shock proteins and stress response-related genes (such as HSPA1B, JUN, FOS, etc.). Typical immune markers are shown in Fig. 3C.

Subsequently, we calculated the RiboSis activity scores for each T cell subtype. The results indicated higher RiboSis activity in CD4_TN, CD4_TFH, CD8_TISG, and CD4_Treg (Fig. 3D). Within the marker genes of

the four cell clusters annotated as CD4_TN (C0, C2, C6, C7, and C17), we identified numerous RiboSis-related genes such as RPS8, RPS6, RPL11, among others. Furthermore, we observed strong intercellular communication between CD4_TN and CD8_Teff_GZMK, as well as between CD4_TN and CD4_Treg (Fig. 3E). To explore the developmental trajectory of CD4⁺ T cells, we conducted pseudotime analysis. The results revealed two differentiation paths, with CD4_TN as a common starting point. The endpoint of trajectory 1 was CD4_TFH and CD4_Treg, while the endpoint of trajectory 2 was CD4_TSTR (Fig. 3F-H, Fig. S1F), unveiling the sequential activation and exhaustion processes of CD4⁺ T cells in tumor microenvironment (TME). The RiboSis activity

scores from the pseudotime analysis indicated a gradual decrease in RiboSis activity as differentiation progressed, with CD4_{TN} exhibiting higher RiboSis activity scores compared to CD4_{TFH}, CD4_{Treg}, and CD4_{TSTR} (Fig. 3I). The expression levels of two hallmark RiboSis-related genes, RPL11 and RPS8, also decreased as differentiation proceeded (Fig. 3J). A heatmap illustrated the expression patterns of genes showing the highest differential expression at various stages of CD4⁺ T cell differentiation, emphasizing the high expression of RiboSis-related genes (RPS8, RPS11, RPS16, etc.) in the early stages of CD4⁺ T cell development (Fig. 3K). Enrichment analysis results highlighted a significant enrichment of protein synthesis-related pathways (response to unfolded protein, protein refolding, etc.) in the early stages of CD4⁺ T cell development (Fig. S1G).

Developing RBS utilizing integrative machine learning techniques

By integrating DEGs between high and low RiboSis-AUC groups and the top 150 genes most strongly correlated with RiboSis, a total of 830 genes were identified (Fig. 4A). Subsequent differential analysis revealed 369 DEGs between LUAD and normal lung tissues ($P < 0.05$, $|\log_2FC| > 0.8$) (Fig. 4B). Univariate Cox regression analysis identified 83 RiboSis-related genes with prognostic significance ($P < 0.05$, Table S2). In the model construction process, the TCGA-LUAD dataset was chosen as the training dataset, while GSE30219, GSE31210, GSE42127, and GSE68465 were selected as validation sets, with batch effects removed (Fig. 4C).

To develop RBS, a total of 81 models were constructed using 10 machine learning algorithms. The average C-index value for each model across all validation datasets was computed. The model built using the CoxBoost and SuperPC algorithms demonstrated the highest average C-index value (0.69) (Fig. 4D). This model encompassed 20 key genes, including AKAP12, ANGPTL4, CA4, CAV1, CCDC68, CD69, CD79A, COL4A3, DAAM2, EFNB2, KIF23, LAMB3, METTL7A, MKI67, PTPRC, RPS19, RRAS, SCNN1B, SLC7A11, and TPX2. Using this algorithm, a risk score was calculated for each patient, subsequently categorizing patients into high-RBS and low-RBS groups in each dataset based on the median risk score.

RBS evaluations

In the aforementioned five datasets, survival analysis was conducted for the high- and low-RBS groups, generating Kaplan-Meier survival curves. The results revealed that the OS of the high-RBS group was significantly lower than that of the low-RBS group ($P < 0.0001$, Fig. 4E-I). In the training dataset (TCGA-LUAD), the high-RBS group exhibited lower PFS rates ($P < 0.05$,

Fig. 4J). Subsequently, time-dependent ROC curves were plotted for RBS in each of the five datasets, and the AUC values were calculated for 1-year, 3-year, and 5-year intervals as follows: 0.7, 0.68, 0.65 (TCGA-LUAD); 0.79, 0.78, 0.76 (GSE30219); 0.68, 0.68, 0.72 (GSE31210); 0.86, 0.74, 0.72 (GSE42127); 0.71, 0.67, 0.62 (GSE68465) (Fig. 4K-O). These findings underscore the robust predictive capabilities of RBS in forecasting the prognosis of LUAD patients. Utilizing PCA analysis for dimensionality reduction and visualization, noticeable disparities in sample distribution exist between the two groups were observed (Fig. 4P, Fig. S2A). Subsequently, we compared the efficacy of RBS and clinical features in predicting the prognosis of LUAD. The risk score of RBS demonstrated a strong correlation with T stage, N stage, and clinical stage in the TCGA-LUAD dataset ($P < 0.05$, Fig. 5A). We plotted the ROC curves related to clinical features, and the results showed that the AUC value for RBS was significantly higher than that for clinical features (Fig. S2B). RBS exhibited the second-highest C-index value across the five datasets, following only clinical stage (Fig. 5B).

Comparison of RBS and previous signatures

In the past few decades, with the rapid advancement of bioinformatics technologies, an increasing number of prognostic signatures have been established to predict the prognosis of LUAD patients [51, 52]. To assess the superiority of RBS, we conducted an extensive PubMed search and gathered a total of 140 published prognostic signatures for LUAD. Subsequently, by calculating their C-index values in aforementioned five datasets, we compared the predictive performance of RBS with other LUAD prognostic signatures. The results revealed that RBS ranked highest in terms of C-index values among four datasets, while placing sixth in the GSE31210 dataset. This implies that RBS surpasses other signatures in predicting the prognosis of LUAD patients (Fig. 5C).

The development and validation of a nomogram

Given the widespread adoption of the TNM staging system in the diagnosis and treatment of LUAD, we integrated RBS with the TNM staging system to construct a comprehensive and detailed nomogram. This integration aims to significantly enhance the predictive performance [53]. Through the utilization of this newly constructed nomogram, we were able to calculate specific scores for each individual patient. These scores allowed for a more precise prognosis assessment, providing a clearer picture of the patient's potential outcome (Fig. S3A). The Cox analysis, utilizing this nomogram, pinpointed age, clinical stage, and the RBS risk score as significant prognostic factors ($P < 0.05$, Fig. S3B). Subsequently, TCGA-LUAD was chosen as the training set and GSE30219 as the test set to assess the precision of this nomogram's predictions.

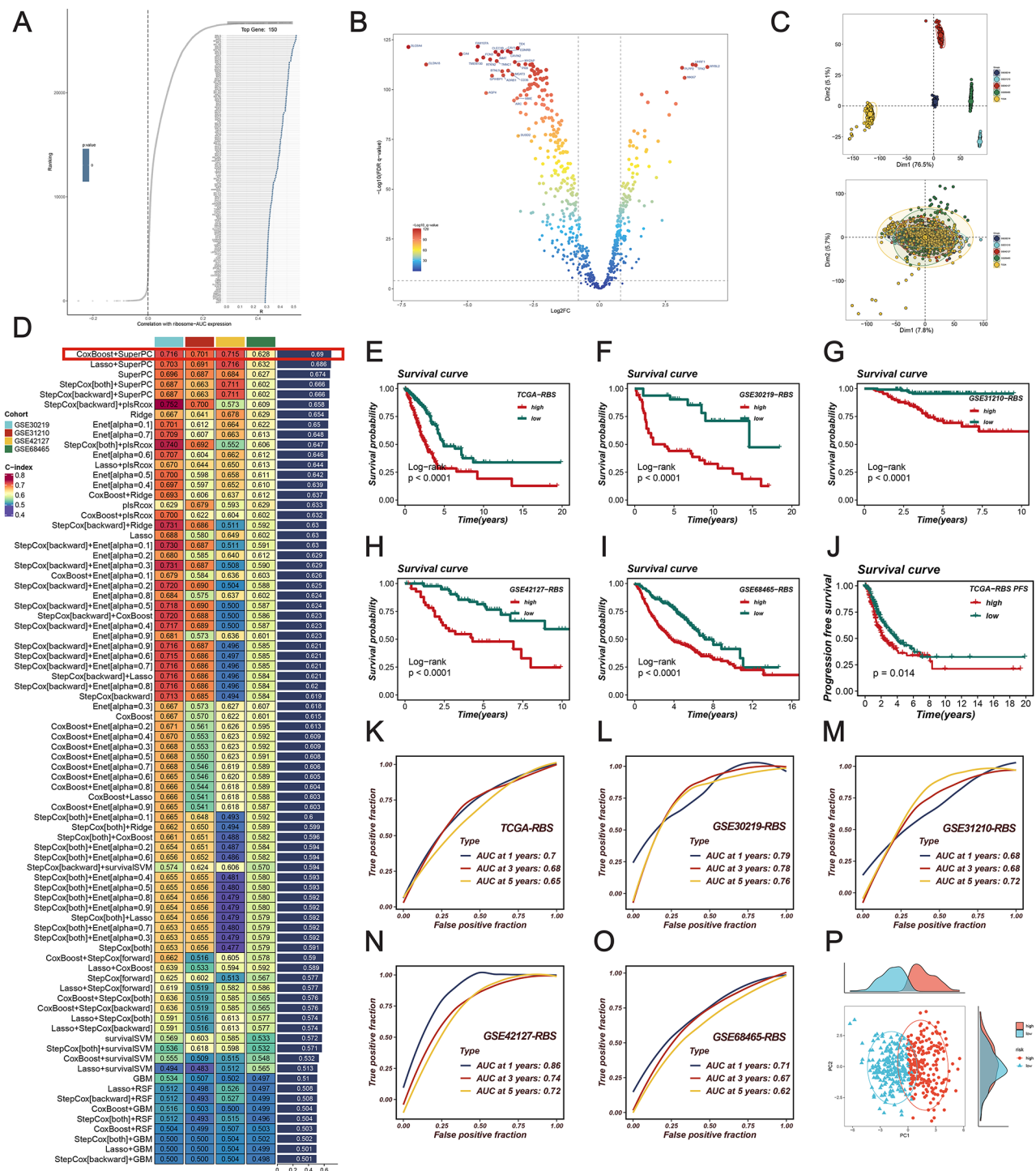


Fig. 4 Construction and Evaluation of RBS. **(A)** Top 150 genes most strongly correlated with RiboSeq. **(B)** Differentially expressed genes (DEGs) between high and low RiboSeq-AUC groups ($P < 0.05$, $|\log_2FC| > 0.7$). **(C)** Sample distribution of five datasets used for model construction before and after batch correction. **(D)** A total of 81 prediction models were established using a machine learning framework, with the C-index values calculated for each model across all validation datasets. Models constructed using CoxBoost and SuperPC algorithms demonstrated the highest average C-index value (0.69). **(E-I)** OS-based K-M curves for high and low RBS groups in TCGA-LUAD ($P < 0.0001$) **(E)**, GSE30219 ($P < 0.0001$) **(F)**, GSE31210 ($P < 0.0001$) **(G)**, GSE42127 ($P < 0.0001$) **(H)**, and GSE68465 ($P < 0.0001$) **(I)**. **(J)** PFS-based K-M curves for high and low RBS groups in TCGA-LUAD ($P < 0.05$). **(K-O)** Time-dependent ROC curves for RBS in TCGA-LUAD **(K)**, GSE30219 **(L)**, GSE31210 **(M)**, GSE42127 **(N)**, and GSE68465 **(O)**. **(P)** Patient distribution of high and low RBS groups visualized in the PCA plot of the TCGA-LUAD dataset

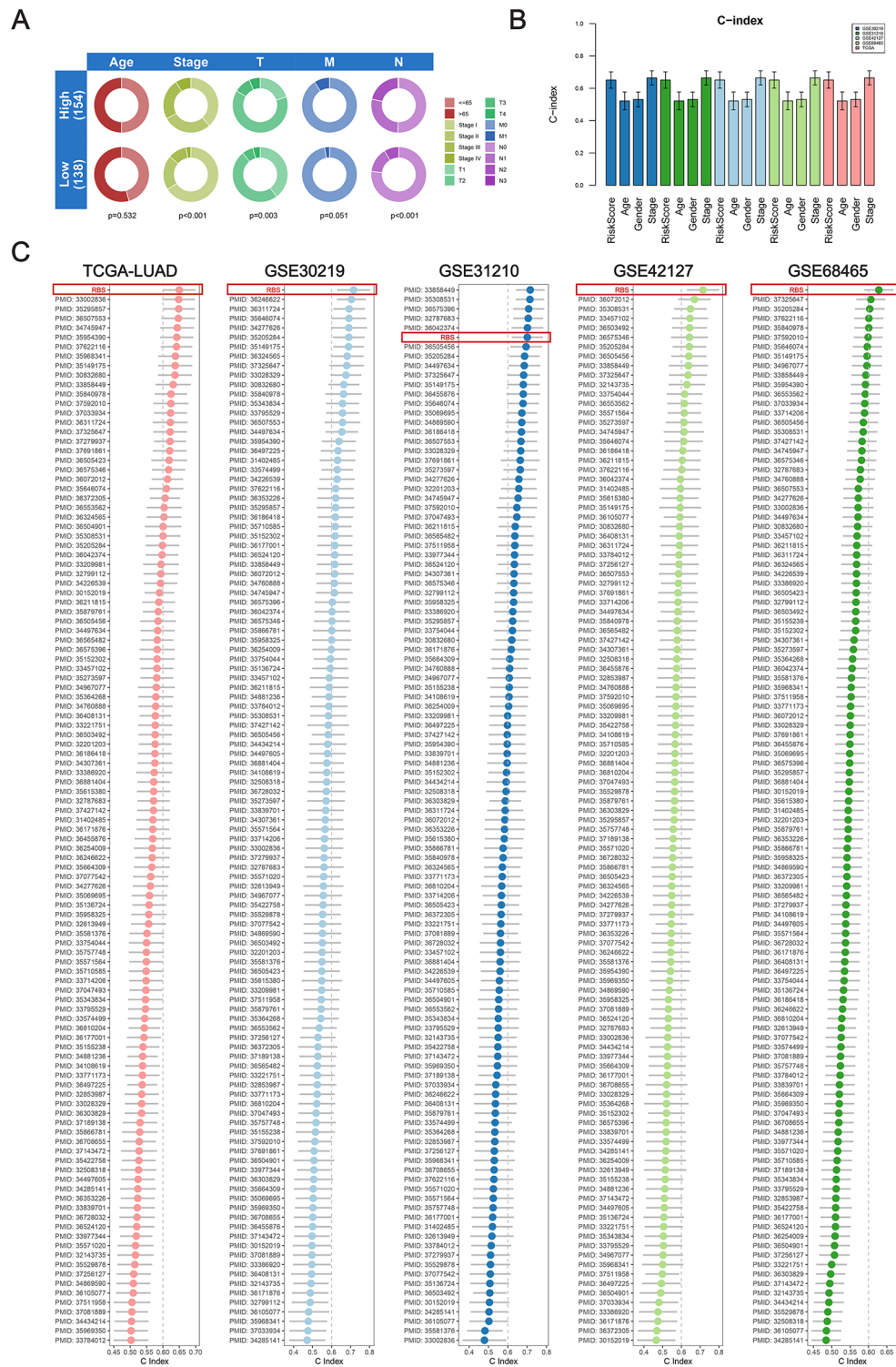


Fig. 5 Evaluation of RBS Prognostic Predictive Ability. **(A)** Correlation of RBS risk score with age, stage, T stage, N stage, and M stage in the TCGA-LUAD dataset. **(B)** Comparison of C-index between RBS risk score and clinical features (age, gender, and stage) in the TCGA, GSE30219, GSE31210, GSE42127, and GSE68465 datasets. **(C)** Calculation and comparison of C-index between RBS and previously published 140 LUAD prognostic models in the TCGA, GSE30219, GSE31210, GSE42127, and GSE68465 datasets

Calibration curves demonstrated concordance between predicted and observed outcomes (Fig. S3C). The AUC values for 1-year, 3-year, and 5-year intervals in the time-dependent ROC curves were as follows: 0.749, 0.749, 0.709 (TCGA-LUAD); 0.538, 0.739, 0.738 (GSE30219) (Fig. S3C). Finally, DCA indicated the potential of this nomogram to accurately predict survival probabilities of LUAD patients at various time intervals (Fig. S3D). These results suggest that this nomogram exhibits strong prognostic predictive performance.

Genomic alterations

Assessing genomic variations and TMB aids in determining molecular subtypes, predicting patient prognosis, and identifying oncogenic driver genes [54, 55]. Fig. S4A illustrates significant chromosomal alterations between high- and low-RBS groups. In the high-RBS group, genes such as TP53, TTN, CSMD3, MUC16, RYR2, and LRP1B exhibit high mutation rates, with a noticeable increase in chromosomal instability (CIN) and more frequent chromosomal deletions or amplifications (Fig. S4B, C). The high-RBS group demonstrates a higher TMB score, with a positive correlation between RBS risk score and TMB score (Fig. S4D, E). To explore the impact of TMB on prognosis, all LUAD patients were divided into high TMB group (H-TMB) and low TMB group (L-TMB) based on TMB scores. Survival analysis indicates that H-TMB group shows a better prognosis (Fig. S4F, $P < 0.05$). Subsequently, survival analysis was conducted by combining TMB scores and RBS risk scores. The results show that the H-TMB + low-RBS group exhibits the best prognosis, while the L-TMB + high-RBS group has the poorest prognosis (Fig. S4G, $P < 0.001$).

Underlying biological mechanisms associated with RBS

The results of GSVA reveal that pathways enriched in the high-RBS group primarily include epithelial-mesenchymal transition (EMT), DNA replication, cell cycle regulation, tricarboxylic acid cycle (TCA cycle), glycolysis, PI3K-AKT-mTOR signaling, and oxidative phosphorylation, among various pro-oncogenic pathways (Fig. 6A, B). Aberrant activation or inhibition of these pathways leads to dysregulated cell cycle control, abnormal DNA replication, and metabolic disturbances, collectively promoting uncontrolled proliferation, genetic instability, enhanced invasion, and metastatic potential of tumor cells, thereby driving rapid progression and deterioration of tumors [56]. This may be a contributing factor to the higher disease risk and poorer prognosis observed in the high-RBS group. Conversely, the low-RBS group demonstrates a favorable prognosis, possibly due to less aggressive tumor biology. This group exhibits lower enrichment levels of pathways associated with unlimited cell proliferation and genetic instability. Metascape analysis indicates

that DEGs between the high- and low-RBS groups are enriched in pathways related to cell cycle regulation and DNA replication (Fig. 6C, D). Furthermore, GSEA reveals that major pathways enriched in the high-RBS group include ribosome and ribosome biogenesis in eukaryotes, while the low-RBS group is enriched in cell adhesion molecules pathways (Fig. 6E).

The correlation between RBS and immune infiltration

The TIME is closely associated with the occurrence and development of tumors [57, 58]. Tumor immunity plays a critical role in the early stages and metastasis of tumor development, including immune surveillance, immune regulation, and tumor immune escape [59]. In order to acquire a more profound understanding of immune-related characteristics of the high-RBS group and low-RBS group, we investigated their potential biological mechanisms. The results obtained demonstrated that the risk score associated with RBS demonstrated an inverse relationship with the immune score, stromal score, and also the ESTIMATE score. Concurrently, it displayed a positive correlation with tumor purity (Fig. S5A). An in-depth analysis utilizing seven distinct algorithms further revealed a notably higher level of immune cell infiltration within the low-RBS group, which encompasses infiltration of CD4⁺ and CD8⁺ T cells (Fig. S5B). ssGSEA results demonstrated that the group with low RBS risk scores showed elevated ssGSEA scores across numerous immune cell types (including activated B cells, activated dendritic cells, eosinophils, macrophages, mast cells, neutrophils, natural killer cells, helper T cells, memory T cells, and memory B cells, among others) (Fig. S5C). Furthermore, the low-RBS group also showed higher ssGSEA scores in immune processes such as antigen-presenting cell (APC) co-stimulation, T cell co-stimulation, type II interferon response, and others (Fig. S5C). Additionally, the expression levels of ICs were increased in the low-RBS group (Fig. S5D), suggesting that patients in the low-RBS group are more likely to benefit from ICB therapy. The risk score of RBS was negatively correlated with key steps in the cancer-immunity cycle, including cancer antigen presentation, recruitment of B cells, CD4⁺ and CD8⁺ T cell recruitment, macrophage recruitment, NK cell recruitment, helper T cell recruitment, Treg cell recruitment, T cell infiltration into the tumor stroma, and killing of cancer cells (Fig. S5E). In the low-RBS group, the levels of immune regulatory factors are elevated, such as MHC molecules HLA-DPA1, HLA-DPB1, and costimulatory molecules such as BTNL8, CD27, SLAMF1 (Fig. S5F). This suggests that a lower RBS risk score may help immune cells enter the TME, enhance their anti-tumor activity, which could be beneficial in immunotherapy. In summary, the low-RBS group exhibits higher immune infiltration and immune regulatory factor levels, creating

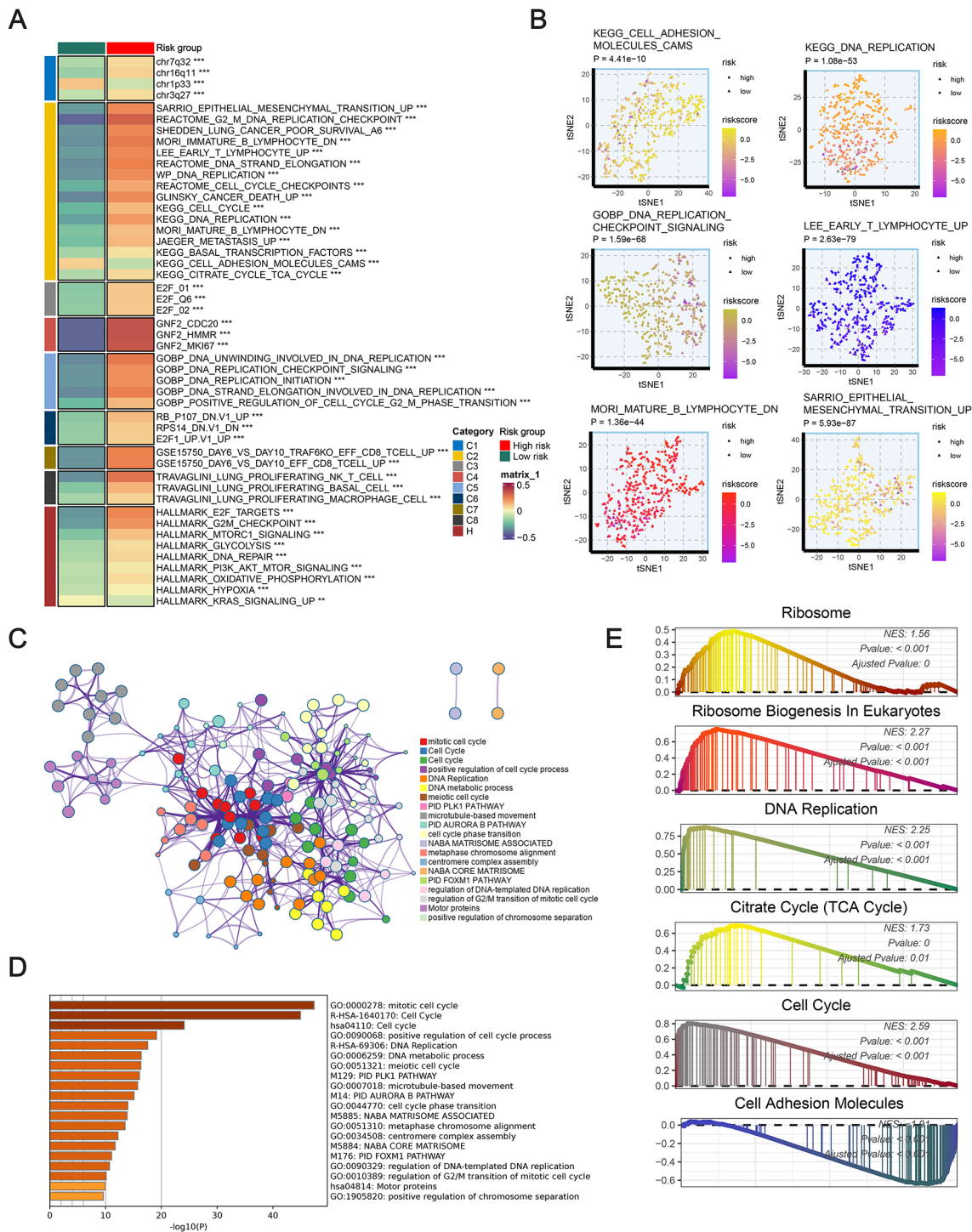


Fig. 6 Underlying Biological Mechanisms Associated with RBS. **(A)** GSEA analysis describing the biological characteristics of high and low RBS groups. **(B)** T-SNE plot revealing differences in pathway activity between high and low RBS groups. **(C, D)** Network plot **(C)** and bar graph **(D)** showing enrichment analysis of differentially expressed genes (DEGs) between high and low RBS groups using Metascape. **(E)** GSEA analysis of gene ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) terms specifically associated with RBS

a TME that is more conducive to effective immune surveillance and response, potentially better controlling tumor growth and spread. This enhanced immune activity may be one of the reasons for the better prognosis of the low-RBS group, as a strong presence of immunity within tumors is frequently linked to better treatment outcomes and extended patient lifespan [59].

Response to immunotherapy and drug sensitivity

Immunotherapy represents a significant breakthrough in cancer treatment in recent years, with notable examples including ICB therapy targeting PD-1 and CTLA-4, as well as chimeric antigen receptor T-cell (CAR-T) therapy [6, 60, 61]. For certain advanced LUAD patients, immunotherapy can help improve their prognosis and extend their survival. Therefore, we further explored the potential role of RBS in predicting immunotherapy response. The results showed that the high-RBS group had higher TIDE scores, indicating a stronger tendency towards immune evasion, which may lead to poorer response to immunotherapy (Fig. 7A, $p < 0.001$). The low-RBS group exhibited higher IPS, suggesting that this group may have a better response to ICB drugs (Fig. 7B, $p < 0.001$). We extrapolated the RBS risk scores for the immunotherapy cohorts by employing the formula and subsequently identified consistent findings across two distinct LUAD immunotherapy cohorts (OAK and POPLAR, Fig. 7C) and three pan-cancer cohorts (three melanoma cohorts, Fig. 7D). The group with lower RBS risk scores exhibited greater suitability for immunotherapy, showing improved prognoses post-treatment (Fig. 7C, D).

Considering the poor response of the high-RBS group to immunotherapy and the fact that in clinical practice, combination therapy involving chemotherapy and targeted therapy persists the primary approach for treating advanced NSCLC, we targeted the high-RBS group to predict potentially effective small molecule drugs using CTRP2.0 and PRISM databases [4, 5, 62]. The aim was to provide new treatment options and improve their prognosis. Ultimately, 24 small molecule drugs were identified (CTRP2.0: 11 drugs, including SB-743921 and paclitaxel; PRISM: 13 drugs, including ispinesib and docetaxel), with lower AUC values in the high-RBS group (Fig. 7E). Paclitaxel, Osimertinib, Gemcitabine, and Sorafenib are four commonly used anticancer drugs. In the PRISM database, the high-RBS group exhibited higher sensitivity to these four drugs (Fig. 7F, $p < 0.001$). The molecular docking of two small molecules, SB-743,921 and ispinesib, with the KIF23 and TPX2 proteins (proteins encoded by two RBS modeling genes) was observed (Fig. 7G, Fig. S6A, B). The binding energies from molecular docking were as follows: -7.3895 kcal/mol (KIF23 + SB-743921), -7.1194 kcal/mol (KIF23 + ispinesib); -5.7081 kcal/mol (TPX2 + SB-743921), -5.9550 kcal/mol (TPX2 + ispinesib). Low binding energies signify a high affinity between the compounds and their protein targets. All binding energies were below -5 kcal/mol, indicating good binding activity between the small molecule drugs and the proteins encoded by the modeling genes. Analysis using the GDSC database revealed significantly higher IC50 values for certain chemotherapy drugs (such as Oxaliplatin, 5-Fluorouracil, Cytarabine, and Cyclophosphamide)

and targeted therapy drugs (such as Gefitinib, Crizotinib, and Sorafenib) in the high-RBS group. This suggests that the high-RBS group may respond better to these drugs, potentially compensating for their poor response to immunotherapy (Fig. S6C).

Immunohistochemistry and in vitro experiments

Previous studies demonstrated CD8 and CD4's correlation with CD8⁺ T cells and CD4⁺ T cells respectively, both capable of anti-tumor activity and predicting favorable immunotherapeutic outcomes [63, 64]. Additionally, heightened expression of the immune checkpoint protein programmed death-ligand 1 (PD-L1) may signify improved immunotherapy efficacy [65]. We collected transcriptome sequencing data from 10 LUAD samples and calculated their respective RBS risk scores. Based on these scores, we classified a sample with the highest RBS scores into the high-RBS group, and a sample with the lowest scores into the low-RBS group. Immunohistochemical analysis revealed that, in contrast to the high RBS group, the low RBS group displayed elevated expressions of CD8, CD4, and PD-L1 (Fig. 8A).

The RBS consists of 20 model genes. Utilizing transcriptome data from TCGA and GTEx databases, we compared the differential expression of these 20 model genes between LUAD and normal lung tissues (Fig. S7A). Using pan-cancer datasets, we analyzed the differential expression of the 20 model genes in tumor and normal tissues at a pan-cancer level (Fig. 8B). IHC images obtained from the Human Protein Atlas (HPA) database further validated the differential expression of these genes at the protein level (Fig. S7B). We evaluated the correlation between the 20 model genes and the RBS using ROC curves and correlation analysis, where KIF23 exhibited a significant positive correlation with RBS (correlation = 0.9, AUC = 0.958, Fig. S7C, Fig. 8C). KIF23, a member of the kinesin family, involved in cell division regulation, is highly expressed in lung cancer, associated with lung cancer prognosis, and may serve as a novel target for lung cancer therapy [66]. Survival analysis indicated a significant negative correlation between KIF23 expression and prognosis of LUAD patients (Fig. 8D). The expression of KIF23 was assessed using qRT-PCR in LUAD cell lines (including A549 and H1299) and human bronchial epithelial cell line 16HBE. Results indicated upregulated expression of KIF23 in LUAD cell lines (Fig. 8E). To investigate the functional role of KIF23 in the development of LUAD, we altered the expression levels of KIF23 in A549 cells by employing specific siRNA for knockdown and overexpression plasmids (Fig. 8F). Results from CCK-8 assays and colony formation experiments suggested that KIF23 may promote the proliferation of LUAD cells (Fig. 8G, H). Subsequently, we investigated the potential impact of KIF23

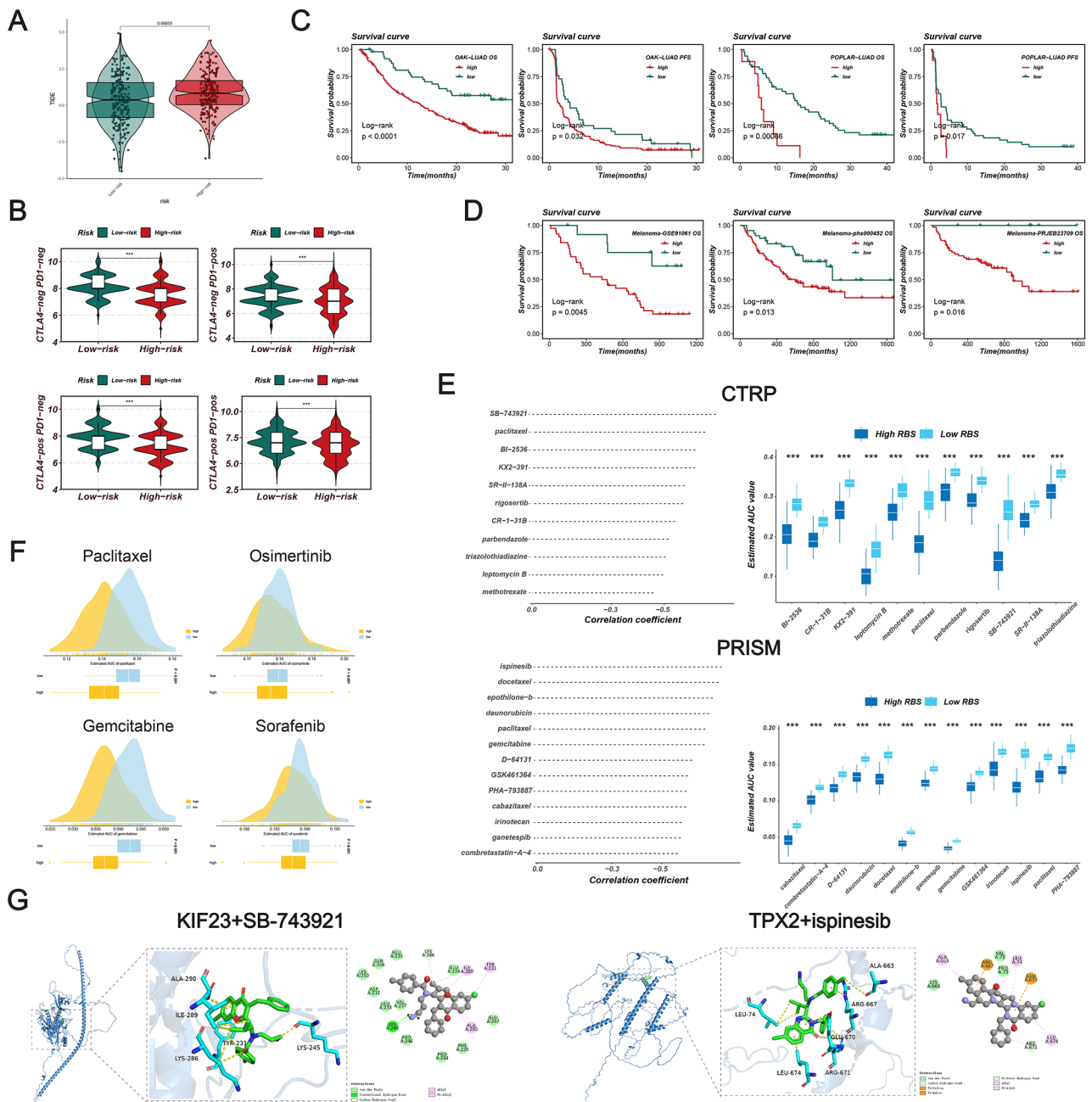


Fig. 7 Response to Immunotherapy and Drug Sensitivity. **(A)** Difference in TIDE scores between high and low RBS groups. **(B)** Prediction of IPS in TCGA-LUAD patients using cancer immunogenomic landscape, consistently indicating higher IPS and greater immunotherapy sensitivity in the low RBS group. **(C)** Calculation of RBS risk scores in two LUAD immunotherapy cohorts (OAK, POPLAR) based on the model formula, followed by survival analysis showing consistent prognostic stratification. **(D)** Evaluation of RBS in three pan-cancer immunotherapy cohorts (Melanoma-GSE91061, Melanoma-phs000452, and Melanoma-PRJEB23709), followed by survival analysis showing consistent prognostic stratification. **(E)** Construction of a drug resistance prediction model based on the CTRP2.0 and PRISM datasets, identifying 11 CTRP derivatives and 13 PRISM derivatives. **(F)** Comparison of sensitivity to four commonly used clinical drugs (Paclitaxel, Osimertinib, Gemcitabine, and Sorafenib) between high and low RBS groups using the PRISM database. **(G)** Molecular docking models illustrating the interaction between the KIF23 protein and SB-743,921, as well as the interaction between the TPX2 protein and ispinisib

on the migration and invasion of LUAD cells. Wound healing experiments showed that knocking down KIF23 significantly hindered wound closure, while overexpression of KIF23 notably promoted wound healing (Fig. 8I). The results of Transwell assays revealed that the silencing

of KIF23 impeded the migration and invasion of LUAD cells, whereas its overexpression markedly boosted their migratory and invasive capabilities (Fig. 8J). These observations validate KIF23's role in stimulating the

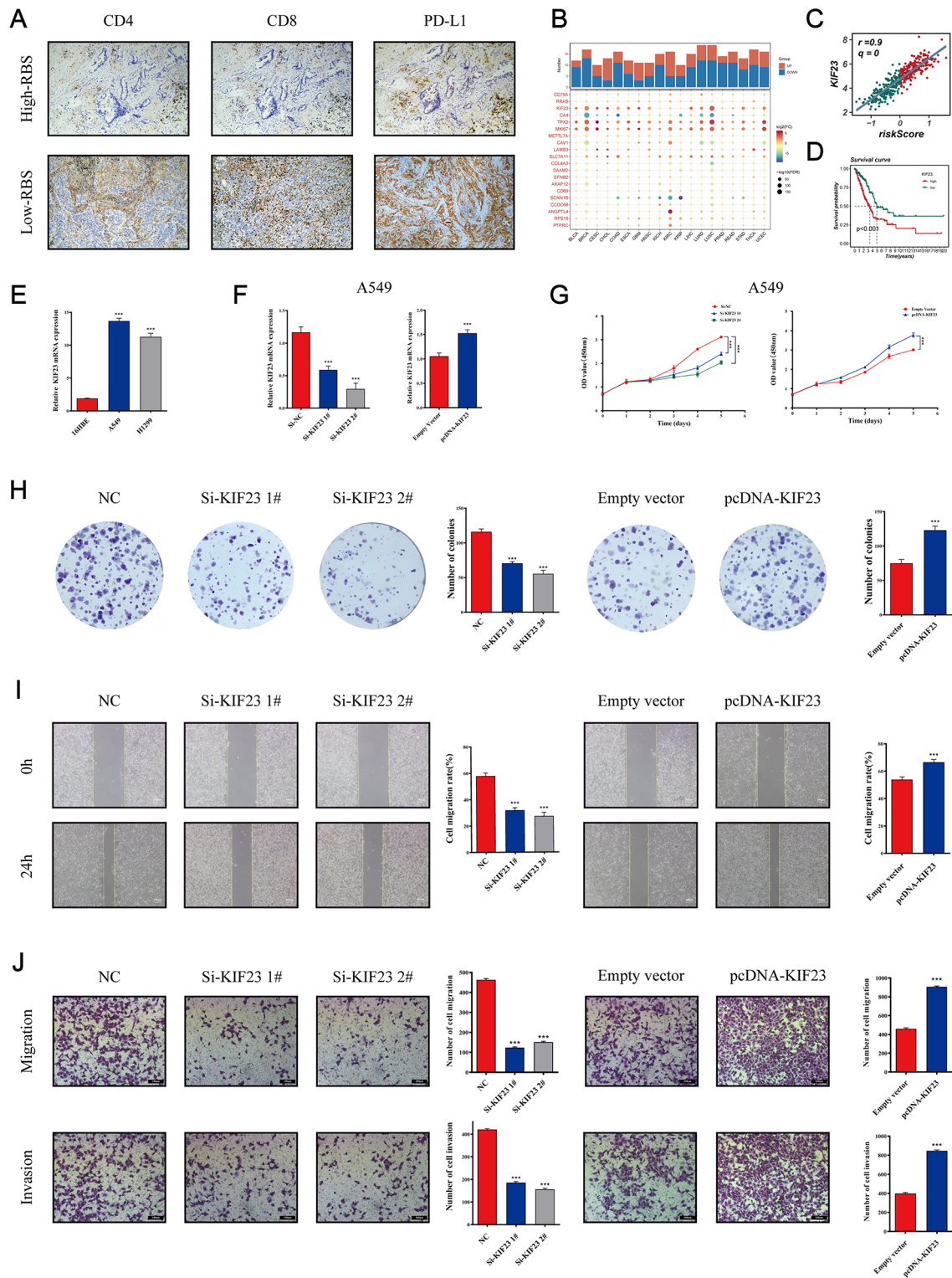


Fig. 8 The Effect of KIF23 on Proliferation, Migration, and Invasion of Lung Adenocarcinoma (LUAD) Cells. **(A)** Validation of the expression differences of CD8, CD4, and PD-L1 between high and low RBS groups through immunohistochemical staining experiments. **(B)** Differential expression of 20 model genes at the pan-cancer level in tumor and normal tissues. **(C)** Correlation analysis demonstrating the relationship between KIF23 and RBS risk score. **(D)** Survival analysis showing a significant negative correlation between the expression level of KIF23 and the prognosis of LUAD patients. **(E)** Evaluation of KIF23 expression using qRT-PCR in LUAD cell lines (including A549 and H1299) and human bronchial epithelial cell line 16HBE. **(F)** Regulation of KIF23 expression in A549 cells using specific siRNA and overexpression plasmids. **(G-J)** Determination of the effects of KIF23 knockdown and overexpression on proliferation of A549 cells using CCK-8 assay **(G)** and colony formation assay **(H)**. Exploration of the effects of KIF23 knockdown and overexpression on migration and invasion of A549 cells through scratch healing assay **(I)** and Transwell assay **(J)**

proliferation, invasion, and migration of A549 cells, suggesting its potential in driving LUAD progression.

Discussion

Lung cancer poses a significant threat to human health, with a persistently high mortality rate, making it a major public health concern [4, 67]. Identifying reliable and effective biomarkers for lung cancer has become an urgent task. Although various lung cancer biomarkers have been discovered using bioinformatics methods, such as the LUAD prognostic model based on B cell marker genes established by Song et al., the reliability and clinical translational potential of these prognostic models remain to be thoroughly investigated due to the lack of model comparisons and clinical validations [68]. In recent years, scRNA-seq technology has been widely employed in cancer research, with researchers aiming to explore the heterogeneity of tumor cells, map pan-cancer single-cell atlases, identify specific cell subpopulations, and investigate their functions within TME [17]. For instance, Tang et al. constructed a pan-cancer single-cell panorama of human natural killer cells and elucidated specific NK cell subtypes [69]. However, these studies tend to focus more on elucidating biological mechanisms and less on exploring the relationship between target cell subpopulations and cancer prognosis and treatment outcomes. RiboSis is a core step in the process of tumor initiation and progression, with considerable evidence suggesting a close association between RiboSis and cancer occurrence, metastasis, and treatment resistance, holding potential clinical translational value as a target for future cancer therapies [9, 12, 14]. Some studies have investigated the role of RiboSis in proliferation and metastasis in LUAD using laboratory techniques [15, 16]. However, there is currently no literature elucidating whether RiboSis-related genes can serve as biomarkers for LUAD and play a role in prognosis prediction and treatment efficacy assessment.

In this study, we first utilized LUAD scRNA-seq data to construct a single-cell atlas, focusing on describing the cell subpopulations of T cells. We identified two relatively novel T cell subtypes: CD8⁺ interferon (IFN) response T cells and CD4⁺ stress response T cells. Both are associated with stress responses and play distinct roles within the TIME. Subsequently, we delved into the association between cell subpopulations and RiboSis, revealing that CD4⁺ naive-like T cells exhibit high expression of RiboSis-related genes, with RiboSis activity gradually decreasing as development progresses. This discovery paves the way for investigating naive-like T cells and RiboSis in LUAD. Building upon LUAD bulk RNA-seq data and RiboSis-related genes, we constructed a prognostic model (RBS). Comparing it with 140 previously published LUAD prognostic models, we confirmed the

superior predictive performance of RBS in prognosis forecasting.

RBS comprises 20 key genes, including AKAP12, ANGPTL4, CA4, CAV1, CCDC68, CD69, CD79A, COL4A3, DAAM2, EFNB2, KIF23, LAMB3, METTL7A, MKI67, PTPRC, RPS19, RRAS, SCNN1B, SLC7A11, and TPX2. AKAP12, a type A kinase anchor protein scaffold, loss of which increases cancer susceptibility, has been reported in studies to be improved in LUAD by miR-338-3p inhibition of AKAP12 [70]. CD69 serves as a leukocyte and NK cell activation marker, playing a crucial role in the pathogenesis of tissue damage under activation of different leukocyte subpopulations and various inflammatory states, with previous research confirming its close association with lung cancer [71]. The gene encoding ribosomal protein S19 (RPS19) is one of the pathogenic genes for Diamond-Blackfan anemia (DBA), with no reported function in lung cancer to date [72, 73]. Targeting protein for Xklp2 (TPX2) is associated with the spindle pole localization of the dynamic protein-dependent Xklp2. TPX2 is highly expressed in various cancers, including lung cancer, with studies confirming its regulation of macrophage polarization in LUAD through the NF- κ B pathway [74].

After validation, RBS has demonstrated strong predictive abilities for the prognosis of LUAD patients. The high-RBS group is associated with multiple pro-oncogenic pathways, showing fewer immune cell infiltrates and a lower response to immunotherapy, potentially contributing to higher disease risk and poorer prognosis. Conversely, the low-RBS group exhibits a TME more conducive to effective immune surveillance and response, displaying a higher response to immunotherapy and better prognosis. At the cellular and molecular biology level, we selected one of the modeling genes, KIF23, for in vitro experiments, confirming its role in promoting proliferation and invasion of LUAD cells. This provides foundational biological evidence for the clinical translation of RBS. KIF23, a member of the kinesin motor protein family involved in cell division regulation, is highly expressed in lung cancer and associated with lung cancer prognosis, potentially serving as a novel target for lung cancer treatment [66]. Previous studies have shown that RNA interference-mediated depletion of KIF23 inhibits in vivo lung tumor formation and induces apoptosis in lung cancer cell lines [75].

The RBS developed in this study is highly associated with the prognosis of LUAD patients, demonstrating superior performance in prognostic prediction and exhibiting significant immune relevance, highlighting its potential as a biomarker in clinical practice. Immunotherapy has been a major breakthrough in cancer treatment in recent years, improving the prognosis and extending the survival of certain advanced LUAD patients [6, 7].

RBS can predict the differential therapeutic responses to immunotherapy among LUAD patients, thereby facilitating the identification of LUAD patient subgroups that are sensitive to immunotherapy. Integrating RBS into future LUAD clinical practice may facilitate personalized risk assessment and immunotherapeutic intervention, ultimately improving patient outcomes and guiding clinical decision-making. However, this study has several limitations. First, it is a retrospective study based on public databases. The limited sample size in the validation datasets may restrict the generalizability of the findings, necessitating further independent validation through robust prospective multicenter cohort studies. Additionally, more LUAD immunotherapy datasets should be incorporated to comprehensively validate the predictive power of RBS in assessing immunotherapy response. Second, the single-cell analysis was not comprehensive, as it primarily focused on T cell subpopulations. Finally, we only conducted in vitro validation experiments for KIF23, leaving the exploration of the biological functions of RiboSis-related genes in LUAD incomplete. Future studies should aim to include a more diverse patient population and validate our findings across multiple independent cohorts to ensure their robustness and applicability. Furthermore, additional experimental investigations, including in vivo studies, are required to elucidate the precise mechanisms by which KIF23 influences LUAD cell proliferation and migration, as well as its downstream signaling pathways.

Conclusion

This study employed scRNA-seq technology to construct a single-cell atlas of LUAD and identified that CD4⁺ naive-like T cells exhibit high expression of RiboSis-related genes. Notably, RiboSis activity gradually decreased as CD4⁺ T cells underwent differentiation and maturation. RiboSis, a fundamental process for ribosome synthesis, is closely associated with tumor initiation, progression, and therapeutic resistance, and holds promise as a potential target for future cancer treatments. Accordingly, we developed a RiboSis-based biomarker—referred to as RBS—by integrating multiple machine learning algorithms, including Lasso, SuperPC, and Cox-Boost. Through a comprehensive multi-omics analytical framework, we systematically investigated the relationship between the RBS risk score and patient prognosis, TIME, response to immunotherapy, and drug sensitivity in LUAD. Compared to 140 existing prognostic biomarkers, RBS demonstrated superior performance in predicting clinical outcomes in LUAD patients. Moreover, RBS exhibited significant immune relevance. The low-RBS subgroup was characterized by a TME more favorable for effective immune surveillance and response, enhanced responsiveness to immunotherapy, and improved overall prognosis. The ability of RBS to stratify LUAD patients

according to differential responses to immunotherapy facilitates the identification of immunotherapy-sensitive subpopulations. Incorporating RBS into future clinical practice may enable personalized risk assessment and immune-based therapeutic strategies, ultimately improving patient outcomes and guiding clinical decision-making. As a core gene within the RBS model, KIF23 was experimentally validated in vitro to promote LUAD cell proliferation, migration, and invasion, supporting its role as a key oncogene in LUAD and highlighting its potential as a therapeutic target. We believe the findings of this study will support future research into the biological mechanisms and clinical translational value of RiboSis in LUAD, offering novel insights and guidance for clinical decision-making.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12967-025-06489-0>.

Supplementary Fig. 1. Processing of scRNA-seq Data. (A) Following batch correction, the cell distribution of 9 LUAD samples is relatively consistent. (B) Proportional plots illustrating the distribution differences of cell types among the 9 LUAD samples. (C) Calculation of RiboSis activity scores for each cell using the “AUCell” package. (D) Statistical analysis of cell communication quantity and cell communication strength between high and low RiboSis-AUC groups. (E) After eliminating batch effects between samples, the distribution of T and NK cells among the 9 LUAD samples is relatively consistent. (F) Pseudotime analysis conducted on CD4⁺ T cells revealing the presence of 5 different states of CD4⁺ T cells. (G) Enrichment analysis based on the pseudotime analysis results of CD4⁺ T cells.

Supplementary Fig. 2. PCA Plot and Clinically Relevant ROC Curves. (A) The PCA plot displays the distribution of patients from the high RBS and low RBS groups in the GSE30219, GSE31210, GSE42127, and GSE68465 datasets. (B) The ROC curves combine clinical features of the TCGA-LUAD, GSE30219, GSE31210, GSE42127, and GSE68465 datasets.

Supplementary Fig. 3. Establishment and Validation of a Nomogram Combined with Clinical Characteristics. (A) In TCGA-LUAD, a nomogram was constructed by integrating RBS, age, and TNM staging to predict the 1, 3, and 5-year survival rates of LUAD patients. (B) A forest plot illustrating age, Stage, and RBS risk score as key prognostic factors in the nomogram. (C) Calibration curves and time-dependent ROC curves of the nomogram were plotted in the training set (TCGA-LUAD). (D) Calibration curves and time-dependent ROC curves of the nomogram were plotted in the testing set (GSE30219). (E) Decision curve analysis (DCA) was performed in the training set (TCGA-LUAD) to further evaluate the prognostic predictive performance of the nomogram.

Supplementary Fig. 4. Multi-omics Characterization of RBS in the TCGA Dataset. (A) Chromosomal amplifications and deletions in high and low RBS groups analyzed using GISTIC 2.0. (B) Genomic alterations in high and low RBS groups. (C) Proportions of genomic alterations in high and low RBS groups. (D) Difference in tumor mutational burden (TMB) between high and low RBS groups. (E) Correlation between TMB and RBS risk score. (F) Kaplan-Meier curves for high TMB group (H-TMB) and low TMB group (L-TMB). (G) Grouping of LUAD patients based on median TMB and RBS risk score for survival analysis.

Supplementary Fig. 5. The Correlation between RBS and Immune Infiltration. (A) Correlation between RBS risk score and immune score, stromal score, ESTIMATE score, and tumor purity. (B) Evaluation of differences in immune cell abundance between high and low RBS groups using the TIMER2.0 database and seven algorithms. (C) Assessment of differences in immune cell abundance and immune function between high and low RBS groups using the ssGSEA algorithm. (D) Differential expression of immune

checkpoint genes between high and low RBS groups. (E) Evaluation of the correlation between cancer immune cycle, immune therapy pathways, and RBS risk score using GSVA. (F) Analysis of immune regulatory gene expression between high and low RBS groups.

Supplementary Fig. 6. Molecular Docking Models and Drug Sensitivity Prediction. (A) Molecular docking model illustrating the interaction between the KIF23 protein and the small molecule drug ispinesib. (B) Molecular docking model illustrating the interaction between the TPX2 protein and the small molecule drug SB-743,921. (C) Comparison of IC50 values of commonly used chemotherapy drugs and targeted drugs between the high RBS group and the low RBS group using the GDSC database.

Supplementary Fig. 7. In-Depth Analysis of 20 Model Genes. (A) Analysis of the differential expression of 20 model genes in LUAD and normal lung tissues using transcriptomic data from the TCGA and GTEx databases. (B) Immunohistochemical (IHC) images from the Human Protein Atlas (HPA) database for 17 modeled genes (CD79A, RRAS, KIF23, CA4, TPX2, MKI67, CAV1, LAMB3, COL4A3, DAAM2, EFN2, AKAP12, CD69, SCNN1B, CCDC68, ANGPTL4, and PTPRC) in normal lung tissues and lung adenocarcinoma tissues (IHC images for METTL7A, RPS19, and SLC7A11 in LUAD are not provided by the HPA database). (C) Correlation analysis between the 20 model genes and RBS using ROC curve analysis.

Supplementary Table 1. The information of RiboSis-related genes.

Supplementary Table 2. 83 RiboSis-related genes with prognostic significance.

Acknowledgements

We thank The Cancer Genome Atlas (TCGA), Gene Expression Omnibus (GEO), UCSC Xena, and Molecular Signature Database (MSigDB) for providing the data used in this study.

Author contributions

ZPS, PPZ, MLZ and LC contributed conception and design of the study; YHW and XG collected the data; ZPS and XCC performed the statistical analysis; ZPS and YHW wrote the first draft of the manuscript; ZHL, JNZ, JWT and LC revised the manuscript; ZPS, PPZ and LC gave the final approval of the version to be submitted. All authors contributed to manuscript and approved the submitted version.

Funding information

This work was supported by the National Natural Science Foundation of China (81972175).

Data Availability

The datasets analysed in the current study are available in the TCGA repository (<https://portal.gdc.cancer.gov/>), GEO (<https://www.ncbi.nlm.nih.gov/geo/>), UCSC Xena database (<https://xenabrowser.net/datapages/>), or the data availability sections of the relevant publications. All data relevant to this investigation, whether generated or analyzed, are comprehensively detailed in this manuscript and its supplementary materials. Further inquiries can be directed to the corresponding authors.

Declarations

Ethics approval and consent to participate

This study was conducted in accordance with the principles of the Declaration of Helsinki. Approval was obtained from the Ethics Committee of the First Affiliated Hospital of Nanjing Medical University (Approval No. 2023-SR-730). Informed consent was obtained from all participants involved in the study.

Conflict of interest

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

Consent to participate

Prior to their participation, all subjects provided their informed consent for inclusion.

Author details

¹Department of Thoracic Surgery, The First Affiliated Hospital of Nanjing Medical University, Nanjing, China

²Department of Oncology, Jiangsu Cancer Hospital, The Affiliated Cancer Hospital of Nanjing Medical University, Jiangsu Institute of Cancer Research, Nanjing, China

³Department of Lung Cancer, Tianjin Lung Cancer Center, Key Laboratory of Cancer Prevention and Therapy, National Clinical Research Center for Cancer, Tianjin's Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin, China

Received: 18 December 2024 / Accepted: 13 April 2025

Published online: 02 May 2025

References

1. Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, Jemal A. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin*. 2024;74(3):229–63.
2. Chen P, Liu Y, Wen Y, Zhou C. Non-small cell lung cancer in China. *Cancer Commun (Lond)*. 2022;42(10):937–70.
3. Howington JA, Blum MG, Chang AC, Balekian AA, Murthy SC. Treatment of stage I and II non-small cell lung cancer: diagnosis and management of lung cancer, 3rd ed: American college of chest physicians evidence-based clinical practice guidelines. *Chest*. 2013;143(5 Suppl):e5278–313.
4. Ettinger DS, Wood DE, Aisner DL, Akerley W, Bauman JR, Bharat A, Bruno DS, Chang JY, Chirieac LR, D'Amico TA, et al. Non-Small cell lung cancer, version 3.2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw*. 2022;20(5):497–530.
5. Gerber DE, Schiller JH. Maintenance chemotherapy for advanced non-small-cell lung cancer: new life for an old idea. *J Clin Oncol*. 2013;31(8):1009–20.
6. Reck M, Remon J, Hellmann MD. First-Line immunotherapy for Non-Small-Cell Lung Cancer. *J Clin Oncol*. 2022;40(6):586–97.
7. Gao Q, Yang L, Lu M, Jin R, Ye H, Ma T. The artificial intelligence and machine learning in lung cancer immunotherapy. *J Hematol Oncol*. 2023;16(1):55.
8. Aghaei M, Khademi R, Bahreiny SS, Saki N. The need to Establish and recognize the field of clinical laboratory science (CLS) as an essential field in advancing clinical goals. *Health Sci Rep*. 2024;7(8):e70008.
9. Elhamamsy AR, Metge BJ, Alsheikh HA, Shevde LA, Samant RS. Ribosome biogenesis: A central player in Cancer metastasis and therapeutic resistance. *Cancer Res*. 2022;82(13):2344–53.
10. Jiao L, Liu Y, Yu XY, Pan X, Zhang Y, Tu J, Song YH, Li Y. Ribosome biogenesis in disease: new players and therapeutic targets. *Signal Transduct Target Ther*. 2023;8(1):15.
11. Ramakrishnan V. Ribosome structure and the mechanism of translation. *Cell*. 2002;108(4):557–72.
12. Pelletier J, Thomas G, Volarević S. Ribosome biogenesis in cancer: new players and therapeutic avenues. *Nat Rev Cancer*. 2018;18(1):51–63.
13. Bursać S, Prodan Y, Pullen N, Bartek J, Volarević S. Dysregulated ribosome biogenesis reveals therapeutic liabilities in Cancer. *Trends Cancer*. 2021;7(1):57–76.
14. Catez F, Dalla Venezia N, Marcel V, Zorbas C, Lafontaine DLJ, Diaz JJ. Ribosome biogenesis: an emerging druggable pathway for cancer therapeutics. *Biochem Pharmacol*. 2019;159:74–81.
15. Wang X, Zhang H, Sapio R, Yang J, Wong J, Zhang X, Guo JY, Pine S, Van Remmen H, Li H, et al. SOD1 regulates ribosome biogenesis in KRAS mutant non-small cell lung cancer. *Nat Commun*. 2021;12(1):2259.
16. Lian Z, Yan X, Diao Y, Cui D, Liu H. T cell differentiation protein 2 facilitates cell proliferation by enhancing mTOR-mediated ribosome biogenesis in non-small cell lung cancer. *Discov Oncol*. 2022;13(1):26.
17. Zhang Y, Wang D, Peng M, Tang L, Ouyang J, Xiong F, Guo C, Tang Y, Zhou Y, Liao Q, et al. Single-cell RNA sequencing in cancer research. *J Exp Clin Cancer Res*. 2021;40(1):81.
18. Lei Y, Tang R, Xu J, Wang W, Zhang B, Liu J, Yu X, Shi S. Applications of single-cell sequencing in cancer research: progress and perspectives. *J Hematol Oncol*. 2021;14(1):91.
19. Seijo LM, Peled N, Ajona D, Boeri M, Field JK, Sozzi G, Pio R, Zulueta JJ, Spira A, Massion PP, et al. Biomarkers in lung Cancer screening: achievements, promises, and challenges. *J Thorac Oncol*. 2019;14(3):343–57.

20. Mino-Kenudson M, Schalper K, Cooper W, Dacic S, Hirsch FR, Jain D, Lopez-Rios F, Tsao MS, Yatabe Y, Beasley MB, et al. Predictive biomarkers for immunotherapy in lung cancer: perspective from the international association for the study of lung Cancer pathology committee. *J Thorac Oncol*. 2022;17(12):1335–54.
21. Zhu J, Fan Y, Xiong Y, Wang W, Chen J, Xia Y, Lei J, Gong L, Sun S, Jiang T. Delineating the dynamic evolution from preneoplasia to invasive lung adenocarcinoma by integrating single-cell RNA sequencing and Spatial transcriptomics. *Exp Mol Med*. 2022;54(11):2060–76.
22. Rousseaux S, Debernardi A, Jacquiau B, Vitte AL, Vesin A, Nagy-Mignotte H, Moro-Sibilot D, Brichon PY, Lantuejoul S, Hainaut P, et al. Ectopic activation of germline and placental genes identifies aggressive metastasis-prone lung cancers. *Sci Transl Med*. 2013;5(186):186ra166.
23. Okayama H, Kohno T, Ishii Y, Shimada Y, Shiraishi K, Iwakawa R, Furuta K, Tsuta K, Shibata T, Yamamoto S, et al. Identification of genes upregulated in ALK-positive and EGFR/KRAS/ALK-negative lung adenocarcinomas. *Cancer Res*. 2012;72(1):100–11.
24. Tang H, Xiao G, Behrens C, Schiller J, Allen J, Chow CW, Suraokar M, Corvalan A, Mao J, White MA, et al. A 12-gene set predicts survival benefits from adjuvant chemotherapy in non-small cell lung cancer patients. *Clin Cancer Res*. 2013;19(6):1577–86.
25. Shedden K, Taylor JM, Enkemann SA, Tsao MS, Yeatman TJ, Gerald WL, Eschrich S, Jurisica I, Giordano TJ, Misk DE, et al. Gene expression-based survival prediction in lung adenocarcinoma: a multi-site, blinded validation study. *Nat Med*. 2008;14(8):822–7.
26. Rittmeyer A, Barlesi F, Waterkamp D, Park K, Ciardiello F, von Pawel J, Gadgeel SM, Hida T, Kowalski DM, Dols MC, et al. Atezolizumab versus docetaxel in patients with previously treated non-small-cell lung cancer (OAK): a phase 3, open-label, multicentre randomised controlled trial. *Lancet*. 2017;389(10066):255–65.
27. Fehrenbacher L, Spira A, Ballinger M, Kowanzet M, Vansteenkiste J, Mazieres J, Park K, Smith D, Ardal-Cortes A, Lewanski C, et al. Atezolizumab versus docetaxel for patients with previously treated non-small-cell lung cancer (POPLAR): a multicentre, open-label, phase 2 randomised controlled trial. *Lancet*. 2016;387(10030):1837–46.
28. Riaz N, Havel JJ, Makarov V, Desrichard A, Urba WJ, Sims JS, Hodi FS, Martín-Algarra S, Mandal R, Sharfman WH, et al. Tumor and microenvironment evolution during immunotherapy with nivolumab. *Cell*. 2017;171(4):934–e949916.
29. Liu D, Schilling B, Liu D, Sucker A, Livingstone E, Jerby-Aron L, Zimmer L, Gutzmer R, Satzger I, Loquai C, et al. Integrative molecular and clinical modeling of clinical outcomes to PD1 Blockade in patients with metastatic melanoma. *Nat Med*. 2019;25(12):1916–27.
30. Gide TN, Quek C, Menzies AM, Tasker AT, Shang P, Holst J, Madore J, Lim SY, Velickovic R, Wongchenko M, et al. Distinct immune cell populations define response to Anti-PD-1 monotherapy and Anti-PD-1/Anti-CTLA-4 combined therapy. *Cancer Cell*. 2019;35(2):238–e255236.
31. Zhang Y, Parmigiani G, Johnson WE. ComBat-seq: batch effect adjustment for RNA-seq count data. *NAR Genom Bioinform*. 2020;2(3):lqaa078.
32. Nerurkar P, Altwater M, Gerhardt S, Schütz S, Fischer U, Weirich C, Panse VG. Eukaryotic ribosome assembly and nuclear export. *Int Rev Cell Mol Biol*. 2015;319:107–40.
33. Zang Y, Ran X, Yuan J, Wu H, Wang Y, Li H, Teng H, Sun Z. Genomic hallmarks and therapeutic targets of ribosome biogenesis in cancer. *Brief Bioinform*. 2024, 25(2).
34. Stuart T, Butler A, Hoffman P, Hafemeister C, Papalexi E, Mauck WM 3rd, Hao Y, Stoeckius M, Smibert P, Satija R. Comprehensive integration of Single-Cell data. *Cell*. 2019;177(7):1888–e19021821.
35. Becht E, McInnes L, Healy J, Dutertre CA, Kwok IWH, Ng LG, Ginhoux F, Newell EW. Dimensionality reduction for visualizing single-cell data using UMAP. *Nat Biotechnol*. 2018.
36. Aibar S, González-Blas CB, Moerman T, Huynh-Thu VA, Imrichova H, Hulselmans G, Rambow F, Marine JC, Geurts P, Aerts J, et al. SCENIC: single-cell regulatory network inference and clustering. *Nat Methods*. 2017;14(11):1083–6.
37. Hao Y, Hao S, Andersen-Nissen E, Mauck WM 3rd, Zheng S, Butler A, Lee MJ, Wilk AJ, Darby C, Zager M, et al. Integrated analysis of multimodal single-cell data. *Cell*. 2021;184(13):3573–e35873529.
38. Jin S, Plikus MV, Nie Q. CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics. *Nat Protoc*. 2024.
39. Qiu X, Hill A, Packer J, Lin D, Ma YA, Trapnell C. Single-cell mRNA quantification and differential analysis with census. *Nat Methods*. 2017;14(3):309–15.
40. Qiu X, Mao Q, Tang Y, Wang L, Chawla R, Pliner HA, Trapnell C. Reversed graph embedding resolves complex single-cell trajectories. *Nat Methods*. 2017;14(10):979–82.
41. Liu Z, Liu L, Weng S, Guo C, Dang Q, Xu H, Wang L, Lu T, Zhang Y, Sun Z, Han X. Machine learning-based integration develops an immune-derived LncRNA signature for improving outcomes in colorectal cancer. *Nat Commun*. 2022;13(1):816.
42. Yuan Q, Sun J, Hong Z, Shang D. Determining a robust prognostic biomarker for 804 patients with pancreatic cancer using a machine learning computational framework. *Int J Surg*. 2025;111(1):1561–3.
43. Ye B, Fan J, Xue L, Zhuang Y, Luo P, Jiang A, Xie J, Li Q, Liang X, Tan J et al. iMLGAM: Integrated Machine Learning and Genetic Algorithm-driven Multomics analysis for pan-cancer immunotherapy response prediction. *iMeta*. n/a(n/a):e70011.
44. Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffer HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res*. 2018;28(11):1747–56.
45. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, Mesirov JP. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*. 2005;102(43):15545–50.
46. Zhou Y, Zhou B, Pache L, Chang M, Khodabakhshi AH, Tanaseichuk O, Benner C, Chanda SK. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun*. 2019;10(1):1523.
47. Li T, Fu J, Zeng Z, Cohen D, Li J, Chen Q, Li B, Liu XS. TIMER2.0 for analysis of tumor-infiltrating immune cells. *Nucleic Acids Res*. 2020;48(W1):W509–14.
48. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, Hoang CD, Diehn M, Alizadeh AA. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015;12(5):453–7.
49. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*. 2017;18(1):220.
50. Chu Y, Dai E, Li Y, Han G, Pei G, Ingram DR, Thakkar K, Qin JJ, Dang M, Le X, et al. Pan-cancer T cell atlas links a cellular stress response state to immunotherapy resistance. *Nat Med*. 2023;29(6):1550–62.
51. Zhang L, Cui Y, Mei J, Zhang Z, Zhang P. Exploring cellular diversity in lung adenocarcinoma epithelium: advancing prognostic methods and immunotherapeutic strategies. *Cell Prolif*. 2024;57(11):e13703.
52. Zhang L, Cui Y, Zhou G, Zhang Z, Zhang P. Leveraging mitochondrial-programmed cell death dynamics to enhance prognostic accuracy and immunotherapy efficacy in lung adenocarcinoma. *J Immunother Cancer*. 2024, 12(10).
53. Detterbeck FC, Woodard GA, Bader AS, Dacic S, Grant MJ, Park HS, Tanoue LT. The proposed ninth edition TNM classification of lung Cancer. *Chest*. 2024;166(4):882–95.
54. Jardim DL, Goodman A, de Melo Gagliato D, Kurzrock R. The challenges of tumor mutational burden as an immunotherapy biomarker. *Cancer Cell*. 2021;39(2):154–73.
55. Palmeri M, Mehnert J, Silk AW, Jabbar SK, Ganesan S, Popli P, Riedlinger G, Stephenson R, de Meritens AB, Leiser A, et al. Real-world application of tumor mutational burden-high (TMB-high) and microsatellite instability (MSI) confirms their utility as immunotherapy biomarkers. *ESMO Open*. 2022;7(1):100336.
56. Rajabi L, Ebrahimdoost M, Mohammadi SA, Soleimani Samarkhazan H, Khamisipour G, Aghaei M. Aqueous and ethanolic extracts of Moringa oleifera leaves induce selective cytotoxicity in Raji and Jurkat cell lines by activating the P21 pathway independent of P53. *Mol Biol Rep*. 2025;52(1):102.
57. Fu T, Dai LJ, Wu SY, Xiao Y, Ma D, Jiang YZ, Shao ZM. Spatial architecture of the immune microenvironment orchestrates tumor immunity and therapeutic response. *J Hematol Oncol*. 2021;14(1):98.
58. Aghapour SA, Torabizadeh M, Bahreini SS, Saki N, Jalali Far MA, Yousefi-Avarand A, Dost Mohammad Ghasemi K, Aghaei M, Abolhasani MM, Sharifani MS, et al. Investigating the dynamic interplay between cellular immunity and tumor cells in the fight against cancer: an updated comprehensive review. *IJC*. 2024;16(2):84–101.
59. Lei X, Lei Y, Li JK, Du WX, Li RG, Yang J, Li J, Li F, Tan HB. Immune cells within the tumor microenvironment: biological functions and roles in cancer immunotherapy. *Cancer Lett*. 2020;470:126–33.
60. Zhou F, Qiao M, Zhou C. The cutting-edge progress of immune-checkpoint Blockade in lung cancer. *Cell Mol Immunol*. 2021;18(2):279–93.
61. Yuan Q, Liu J, Wang X, Du C, Zhang Y, Lin L, Wang C, Hong Z. Deciphering the impact of dietary habits and behavioral patterns on colorectal cancer. *Int J Surg*. 2025.

62. Miller M, Hanna N. Advances in systemic therapy for non-small cell lung cancer. *BMJ*. 2021;375:n2363.
63. Giles JR, Globig AM, Kaech SM, Wherry EJ. CD8(+) T cells in the cancer-immunity cycle. *Immunity*. 2023;56(10):2231–53.
64. Oh DY, Fong L. Cytotoxic CD4(+) T cells in cancer: expanding the immune effector toolbox. *Immunity*. 2021;54(12):2701–11.
65. Topalian SL, Taube JM, Pardoll DM. Neoadjuvant checkpoint Blockade for cancer immunotherapy. *Science* 2020, 367(6477).
66. Kato T, Wada H, Patel P, Hu HP, Lee D, Ujiie H, Hirohashi K, Nakajima T, Sato M, Kaji M, et al. Overexpression of KIF23 predicts clinical outcome in primary lung cancer patients. *Lung Cancer*. 2016;92:53–61.
67. Saki N, Haybar H, Aghaei M. Subject: motivation can be suppressed, but scientific ability cannot and should not be ignored. *J Transl Med*. 2023;21(1):520.
68. Song P, Li W, Wu X, Qian Z, Ying J, Gao S, He J. Integrated analysis of single-cell and bulk RNA-sequencing identifies a signature based on B cell marker genes to predict prognosis and immunotherapy response in lung adenocarcinoma. *Cancer Immunol Immunother*. 2022;71(10):2341–54.
69. Tang F, Li J, Qi L, Liu D, Bo Y, Qin S, Miao Y, Yu K, Hou W, Li J, et al. A pan-cancer single-cell panorama of human natural killer cells. *Cell*. 2023;186(19):4235–e42514220.
70. Chang J, Liu S, Li B, Huo Z, Wang X, Zhang H. MiR-338-3p improved lung adenocarcinoma by AKAP12 suppression. *Arch Med Sci*. 2021;17(2):462–73.
71. Zhou X, Zhang P, Luo W, Zhang L, Hu R, Sun Y, Jiang H. Ketamine induces apoptosis in lung adenocarcinoma cells by regulating the expression of CD69. *Cancer Med*. 2018;7(3):788–95.
72. Kuramitsu M, Hamaguchi I, Takuo M, Masumi A, Momose H, Takizawa K, Mochizuki M, Naito S, Yamaguchi K. Deficient RPS19 protein production induces cell cycle arrest in erythroid progenitor cells. *Br J Haematol*. 2008;140(3):348–59.
73. Flygare J, Aspesi A, Bailey JC, Miyake K, Caffrey JM, Karlsson S, Ellis SR. Human RPS19, the gene mutated in Diamond-Blackfan anemia, encodes a ribosomal protein required for the maturation of 40S ribosomal subunits. *Blood*. 2007;109(3):980–6.
74. Wang L, Zhang H, Li Y, Li L. TPX2 influences the regulation of macrophage polarization via the NF- κ B pathway in lung adenocarcinoma. *Life Sci*. 2024;340:122437.
75. Iltzsche F, Simon K, Stopp S, Pattschull G, Francke S, Wolter P, Hauser S, Murphy DJ, Garcia P, Rosenwald A, Gaubatz S. An important role for Myb-MuvB and its target gene KIF23 in a mouse model of lung adenocarcinoma. *Oncogene*. 2017;36(1):110–21.

Publisher's note

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