Cancer Cell International

RESEARCH

Open Access

Predictive value of dendritic cell-related genes for prognosis and immunotherapy response in lung adenocarcinoma

Zihao Sun[1,2], Mengfei Hu[3], Xiaoning Huang[1,2], Minghan Song[1,2], Xiujing Chen[1,2], Jiaxin Bei[1,2*], Yiguang Lin[1,5*] and Size Chen[1,2,4,6*]

Abstract

Background Patients with lung adenocarcinoma (LUAD) receiving drug treatment often have an unpredictive response and there is a lack of effective methods to predict treatment outcome for patients. Dendritic cells (DCs) play a significant role in the tumor microenvironment and the DCs-related gene signature may be used to predict treatment outcome. Here, we screened for DC-related genes to construct a prognostic signature to predict prognosis and response to immunotherapy in LUAD patients.

Methods DC-related biological functions and genes were identified using single-cell RNA sequencing (scRNA-seq) and bulk RNA sequencing. DCs-related gene signature (DCRGS) was constructed using integrated machine learning algorithms. Expression of key genes in clinical samples was examined by real-time q-PCR. Performance of the prognostic model, DCRGS, for the prognostic evaluation, was assessed using a multiple time-dependent receiver operating characteristic (ROC) curve, the R package, "timeROC", and validated using GEO datasets.

Results Analysis of scRNA-seq data showed that there is a significant upregulation of LGALS9 expression in DCs isolated from malignant pleural effusion samples. Leveraging the Coxboost and random survival forest combination algorithm, we filtered out six DC-related genes on which a prognostic prediction model, DCRGS, was established. A high predictive capability nomogram was constructed by combining DCRGS with clinical features. We found that patients with a high-DCRGS score had immunosuppression, activated tumor-associated pathways, and elevated somatic mutational load and copy number variant load. In contrast, patients in the low-DCRGS subgroup were resistant to chemotherapy but sensitive to the CTLA-4 immune checkpoint inhibitor and targeted therapy.

Conclusion We have innovatively established a deep learning-based prediction model, DCRGS, for the prediction of the prognosis of patients with LUAD. The model possesses a strong prognostic prediction performance with high accuracy and sensitivity and could be clinically useful to guide the management of LUAD. Furthermore, the findings

*Correspondence:
Jiaxin Bei
beijiaxinya@163.com
Yiguang Lin
yiguang.lin@hotmail.com
Size Chen
chensize@gdpu.edu.cn

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

of this study could provide an important reference for individualized clinical treatment and prognostic prediction of patients with LUAD.

Keywords Machine learning, Prognosis prediction, Lung adenocarcinoma, Dendritic cells, Immunotherapy

Background

Lung cancer is one of the most prevalent cancers with a high fatality rate [1]. Nearly 84% of lung cancers are non-small cell lung cancer (NSCLC), and lung adenocarcinoma (LUAD) is the most common subtype of NSCLC. Immunotherapies have been rapidly developed, but their successful application is severely constrained by the fact that only a minority of LUAD patients can benefit long-term from immunotherapy. Several biomarkers, including PD-L1 expression and tumor mutational burden (TMB), which are frequently employed for immunotherapy response prediction, can only partially characterize the heterogeneity of the tumor microenvironment (TME) [2, 3]. Accurate evaluation of prognosis and improvement of survival in patients with LUAD remains a great challenge. Therefore, it is particularly important to develop a new prognostic signature that can be utilized to accurately predict the response to treatment and the prognosis of patients with LUAD.

Dendritic cells (DC) are the best antigen presenting cells (APC), and DC-based vaccines have been widely used in cancer immunotherapy [4]. DCs can be broadly defined as conventional type I dendritic cells (cDC1s) and conventional type II dendritic cells (cDC2s), both of which play a crucial role in regulating the immune activation of CD8 + T cells and tumor antigen tolerance [5]. Although cancer cells can directly present tumor antigens via their own major histocompatibility complex class I (MHC-I) molecules, cross-presentation of DCs, dedicated antigen-presenting cells, is necessary to maintain the cytotoxic immune response of primitive CD8 + T cells [6]. The elevated abundance of mature dendritic cells (mDCs) infiltrating the tumor microenvironment (TME) indicates a high level of CD8 + T-cell infiltration in lung cancer, which is associated with the long-term survival of NSCLC patients [7, 8]. Stimulatory dendritic cells (sDCs) also play a vital role in stimulating cytotoxic T cells and driving anticancer immunity [9]. Furthermore, activation of plasmacytoid dendritic cells (pDC) in the TME effectively enhances the ability of NK and T cells to recognize and kill tumor cells [10, 11]. Today, transcriptome sequencing using clinical specimens has become an important tool for studying TME [12, 13]. Unfortunately, tumor tissue from LUAD patients is not always available. In terms of clinical diagnosis and the detection of genomic profiling, several investigations have established that malignant pleural effusion (MPE) and lung adenocarcinoma tissue (LAT) are clinically equivalent [14, 15]. MPE can be extensively used as an accessible specimen when tumor tissue is unavailable.

With the successful completion of human genome engineering [16], precision medicine is gradually replacing the traditional medical model [17]; thus, there are now higher requirements for accurate prognosis prediction for cancer patients. Machine learning, which increases the accuracy of models through algorithm iteration, is widely used for the construction of clinical prognostic models. DCs, which enhance the efficacy of immune checkpoint inhibitors (ICIs) by interacting with T cells [18, 19], have become a new breakthrough in tumor immunotherapy. Therefore, we used single-cell RNA sequencing (scRNA-seq) data to explore the functional differences of DCs in MPE and LAT samples and also assessed bulk RNA sequencing (bulk-seq) data to screen DC-associated genes. Machine learning algorithms were used to construct a DC-related gene signature (DCRGS) to predict prognosis, immunotherapy response, and drug screening in patients with LUAD.

Methods

Data source and pre-processing

Our research protocol is depicted in Fig. 1. The gene expression data of LAT used in this study were obtained from The Cancer Genome Atlas (TCGA) and Gene Expression Omnibus (GEO). Among them, five datasets (TCGA-LUAD, GSE26939, GSE31210, GSE42127, and GSE72094) embracing comprehensive overall survival (OS) information were employed to build and validate DCRGS. The scRNA-seq data included five LAT samples and five MPE samples from LUAD patients from GEO and one MPE sample from the First Affiliated Hospital of Guangdong Pharmaceutical University (FAHGPU). The 12 LAT samples used for the real-time quantitative PCR (RTqPCR) assay were obtained from FAHGPU. The reference gene sets used for the gene set enrichment analysis (GSEA) were derived from the Molecular Signatures Database (MSigDB) and previously published studies (Additional file 1: Table S1). Immunotherapy sensitivity data were downloaded from The Cancer Immunome Atlas (TCIA).

Analysis of the scRNA-seq data

We used the R package "Seurat" to quality control the scRNA-seq data from 5 LAT samples and 6 MPE samples. Cells with a total number of genes detected greater than 500, a mitochondrial gene expression ratio less than 20%, an erythrocyte gene expression ratio less

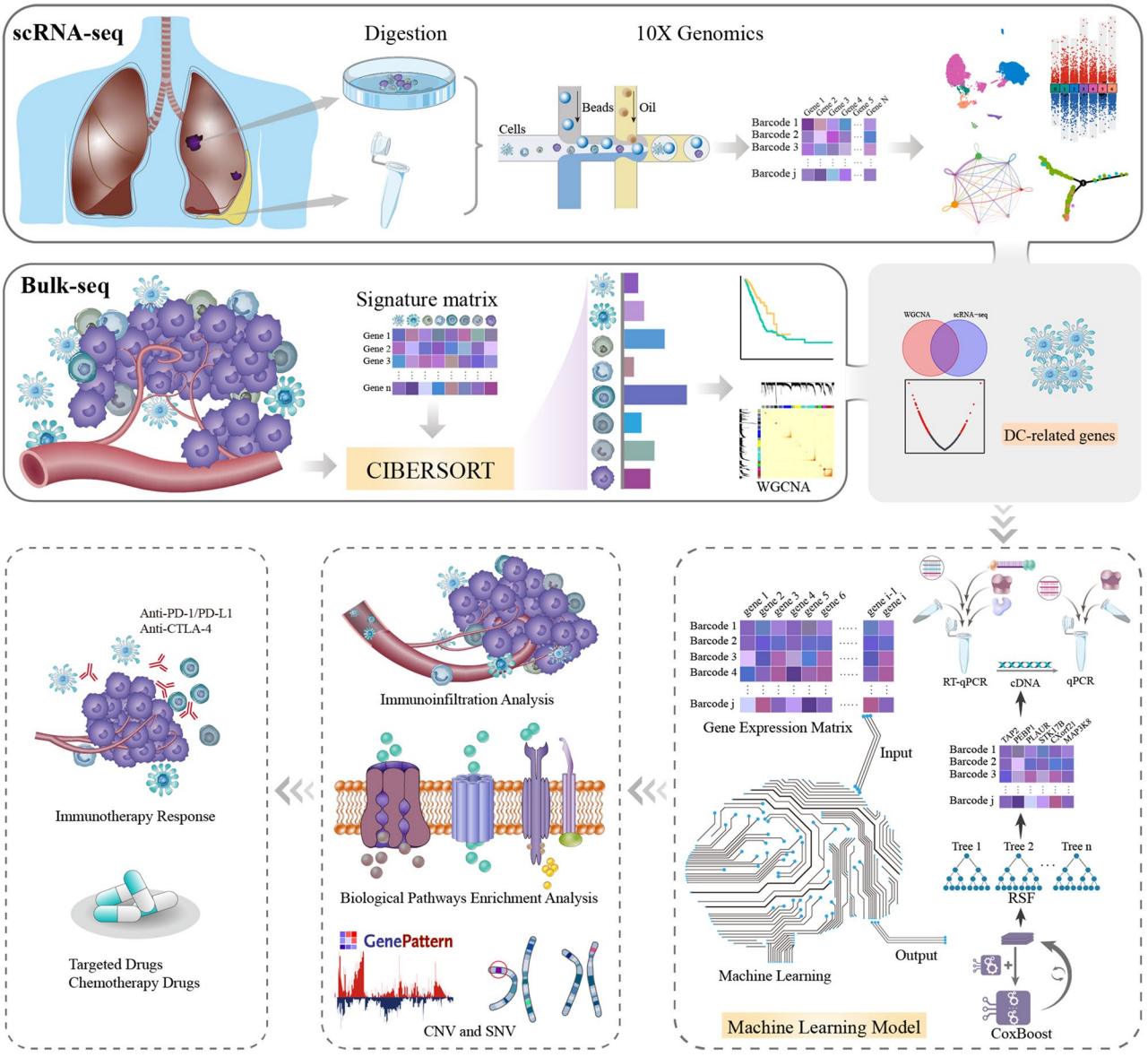

Fig. 1 Flow chart outlining key comprehensive analyses used in the study

than 3%, and a total number of unique molecular identifiers (UMIs) greater than 1000 were retained. Next, the "NormalizeData" and "FindVariableFeatures" functions were utilized to normalize the data and find highly variable genes with nfeatures = 2000. After further processing of the scRNA-seq data by the "ScaleData" and "RunPCA" functions, the R package "harmony" was employed to integrate the data to remove batch effects between samples. Cells were clustered with resolution set to 0.5 based on the "FindNeighbors", "FindClusters" and "RunUMAP" functions. We annotated the cell clusters of the LAT and MPE samples using the marker genes. The "FindAllMarkers" function was used to identify differentially expressed genes for each cell type in the LAT and MPE samples, and then DC-related genes were filtered based on a differential expression multiple of 0.3 and a p value less than 0.05. In addition, we used the R packages "CellChat" and "monocle" for cell–cell interaction analysis and pseudotime cell trajectory analysis, respectively. The thresholds used in the preprocessing of single-cell sequencing data were determined based on multiple references and comprehensive consideration of data quality [20–22].

Weighted gene coexpression network analysis (WGCNA)

TIMER2.0 was applied to infer tumor immunoinfiltration abundance [23], and the optimal cutoff value for DC infiltration abundance was determined by the "surv_cutpoint" function to stratify the TCGA-LUAD cohort. We used the level of DC infiltration abundance as a phenotypic trait and performed WGCNA [24] on

the TCGA-LUAD transcriptome data to identify DC-associated genes. Samples were first clustered using the "hclust" function, and the gene expression matrix was devoted to calculating a weighted network by expression similarity and determining a soft threshold for the network. Then, the network neighbor distance of the gene expression matrix was calculated. The "TOMsimilarity" function was employed to calculate the topological overlap matrix (TOM). The modules were then identified by hierarchical clustering, setting the minimum number of genes in the module at 50. The "moduleEigengenes" function calculates the eigengenes of the module and clusters all modules hierarchically. The "moduleEigengenes" function was performed to calculate the eigengenes of the module, which can be applied to distinguish important modules related to the high or low infiltration abundance of DCs. 2000 genes were randomly selected to map the heatmap of the coexpression network, and the TOM dissimilarity was exponentiated to better show the moderate intensity relationships of gene coexpression. Finally, the module and trait correlation heat maps were mapped by the labeledHeatmap function, and the genes of the modules significantly associated with DCs were extracted.

Derivation of prognostic signature with machine learning algorithms

DC-related genes were further screened by survival analysis, and the expression data of these genes were extracted from the TCGA-LUAD cohort as the training set. Meanwhile, the GSE26939, GSE31210, GSE42127 and GSE72094 cohorts were employed as the validation set. To construct a prognostic signature with superior performance, we used a combination of 101 algorithms consisting of 10 different machine learning algorithms [25], which included random survival forest (RSF), least absolute shrinkage and selection operator (LASSO), ridge, elastic network (Enet), CoxBoost, partial least squares regression for Cox (plsRcox) [26], supervised principal components (SuperPC), gradient boosting machine (GBM), survival support vector machine (Survival-SVM), and stepwise Cox regression (StepCox). Machine learning models were cross-validated using leave-one-out cross-validation (LOOCV). The concordance index (C-index) was used to assess each model's performance, and the best model was then employed to filter the signature genes that were applied to create the DCRGS. The TCGA-LUAD cohort was then divided into low- and high-DCRGS subgroups according to the median DCRGS score, a commonly used threshold. Four GEO-LUAD cohorts were categorized into two subgroups using the same cutoff value as the TCGA-LUAD cohort to demonstrate the enhanced stability and robustness of the prognostic model (Additional file 1: Table S2).

Evaluation of model performance and building nomogram

To further evaluate the performance of the prognostic model, multiple time-dependent receiver operating characteristic (ROC) curve analysis was executed using the R package "timeROC" to assess the performance of DCRGS for the prognostic evaluation of patients aged 1–5 years. Additional clinical characteristics were collected from each cohort and contrasted with the prognostic accuracy of DCRGS to demonstrate its superiority. Furthermore, we retrieved 85 LUAD prognostic features from PubMed and determined the C-index of each feature across five cohorts (Additional file 1: Table S3). Then, we incorporated the DCRGS into a nomogram and assessed its accuracy using decision analysis, the C-index, and a calibration curve. An online intersectional nomogram tool was developed using the R packages "shinydashboard" and "DynNom" to facilitate the clinical application of the prediction model.

RTqPCR

All clinical specimens were retrospectively collected from the First Affiliated Hospital of Guangdong Pharmaceutical University. The study was approved by the Human Ethics Committee of the First Affiliated Hospital of Guangdong Pharmaceutical University (No. 69, 2022). RTqPCR was used to detect the expression levels of six key genes in the LAT samples. Primers for six key genes were designed according to the gene sequences in NCBI (Additional file 1: Table S4), and RTqPCR assays were performed using cDNA as template. In this study, NanoDrop 2000 C was used to determine the level of RNA in the recruited LAT samples. Next, reverse transcription kits were used to create cDNA amplification templates. Then, primers, templates, and samples were blended and kept in a constant-temperature bath at 50 °C for 30 min. Subsequently, the samples were transferred to an 85 °C thermostatic bath for 5 min, followed by cooling on ice. The materials were then divided among three 96-well plates to conduct the experiment.

Annotation of biological characteristics

Tumor immune cell infiltration scores were calculated for each LUAD patient by five algorithms, including TIMER2.0 (https://cistrome.shinyapps.io/timer/), MCP-counter [27], xCell [28], single-sample gene set enrichment analysis (ssGSEA) and ESTIMATE [29]. To further clarify the differences in biological functions in the TME between the two risk subgroups, GSEA was conducted using 15 previously published pathways associated with tumor development and 50 hallmark pathways. Gene Ontology (GO) enrichment analysis of differentially expressed genes from both subgroups was executed in this study. In addition, the expression differences of 74

immune-related regulatory factors in both subgroups were explored.

Integrating analysis of single-nucleotide variant (SNV) and copy number variant (CNV)

SNV and CNV data of the LUAD cohort were obtained from the official TCGA website. The "maftools" package (version 2.10.05) was applied to display the top 20 gene mutation maps and 18 differentially expressed genes ($p < 0.001$) for the two subgroups. The "clusterProfiler" package was used to detect the enrichment of differentially mutated genes in both subgroups in the reference gene set "c2.all.v2023.1.Hs.entrez.gmt" (https://www.gsea-msigdb.org/). Moreover, we used Gistic2.0 (https://cloud.genepattern.org/) to analyze the CNV data online.

Immunotherapy sensitivity analysis and drug screening

We obtained clinical medication information from the TCIA website for the TCGA-LUAD cohort to predict the clinical response to ICIs in patients between the two subgroups. Furthermore, the R package "oncoPredict" was applied to predict the half maximal inhibitory concentration (IC50) of various targeted drugs and small molecule compounds in LUAD patients. A lower IC50 suggests that the drug is more successful in tumor treatment.

Statistical analysis

R software (version 4.1.3) was utilized for statistical analysis and visualization. The data were tested for normality using the "shapiro.test" function. The hypergeometric distribution was utilized to test the significance of the intersection in both sets. RTqPCR data were compared between two groups using the unpaired t-test, while for non-normally distributed data, the Wilcoxon test was performed to compare the statistical differences between the two groups. Spearman's rank test was employed to test the correlation between the two groups of non-normally distributed data. The log-rank test was executed to examine the statistical differences in the results of survival analysis. A two-tailed p-value less than 0.05 was considered a statistically significant difference.

Results

Identification of DCs and related genes by scRNA-seq

The eleven samples were integrated using the Harmony algorithm, and the UMAP clustering of the eleven samples is shown in Fig. 2A. The results indicated that the batch effect between the samples had been minimized. We used marker genes to annotate cell clusters (Fig. 2B). Seven cell types were specified in LAT samples; these included endothelial cells, cDCs, epithelial cells, fibroblasts, macrophages, T cells, and mast cells. Nine cell types were identified in MPE samples; these included B cells, cDCs, epithelial cells, pDCs, fibroblasts, mono/macrophages, neutrophils, plasma cells, and T cells. Figure 2C and D display the differentially expressed genes and proportions of each cell population. And a total of 1448 DC-associated genes were identified from the two groups of samples ($p < 0.05$). We performed a pseudotime investigation on myeloid cells to better comprehend the status of DCs. The findings demonstrated that there were two nodes in the myeloid development trajectory in the LAT samples, and DCs first appeared at a later stage of development (Fig. 2E), while there was only one node in the MPE samples, and neutrophils, mono/macrophages and DCs all appeared at the same time (Fig. 2F).

Annotation of the biological functions of DCs and screening of DC-related genes

Cells in multicellular organisms primarily conduct biological functions through cytokine-mediated intercellular interactions. In LAT samples, DCs predominantly interacted with macrophages as signal transmitters (Fig. 3A), whereas in MPE samples, they strongly connected with both pDCs and macrophages (Fig. 3B). The interaction signals that were markedly activated between these cells were mainly associated with antigen presentation (Fig. 3C). We also found that LGALS9-CD44/CD45 signaling was dramatically enhanced between DCs and monocytes/macrophages in MPE samples compared to LAT samples. One of the reasons for this phenomenon is the higher expression of the LGALS9 gene in the DCs of the MPE samples (Fig. 3D). Further analysis elucidated that the expression levels of HLA-D region genes were markedly elevated in DCs in the LAT and MPE samples (Fig. 3E).

Next, we established 0.56 as the optimal cutoff value for DC infiltration abundance in samples from the TCGA-LUAD cohort (Additional file 1: Fig. S1A). A gene coexpression network of LUAD patients was built by the R package "WGCNA". The soft threshold value of the adjacency matrix was set to 5, and the β value was selected as 10 (Additional file 1: Fig. S1B). After transforming the gene expression matrix of the TCGA-LUAD cohort into adjacency and topology matrices, genes were clustered based on the TOM matrix. Then, the modules with high similarity were merged by setting the shear height to 0.3, resulting in a total of 15 modules (Additional file 1: Fig S1C, D). There was a clear coexpression pattern among the genes of the brown module (Additional file 1: Fig. S1E), which were also highly correlated with the infiltration abundance of DCs (Additional file 1: Fig. S1F). After identifying 2100 DC-related genes using WGCNA, these genes intersected with the 1448 DC-related genes identified through the single-cell annotation of the LAT and MPE samples. As a result, a total of 454 DC-related genes were obtained ($p < 0.001$; Additional file 1: Fig. S1G). Subsequent univariate Cox regression analysis was applied

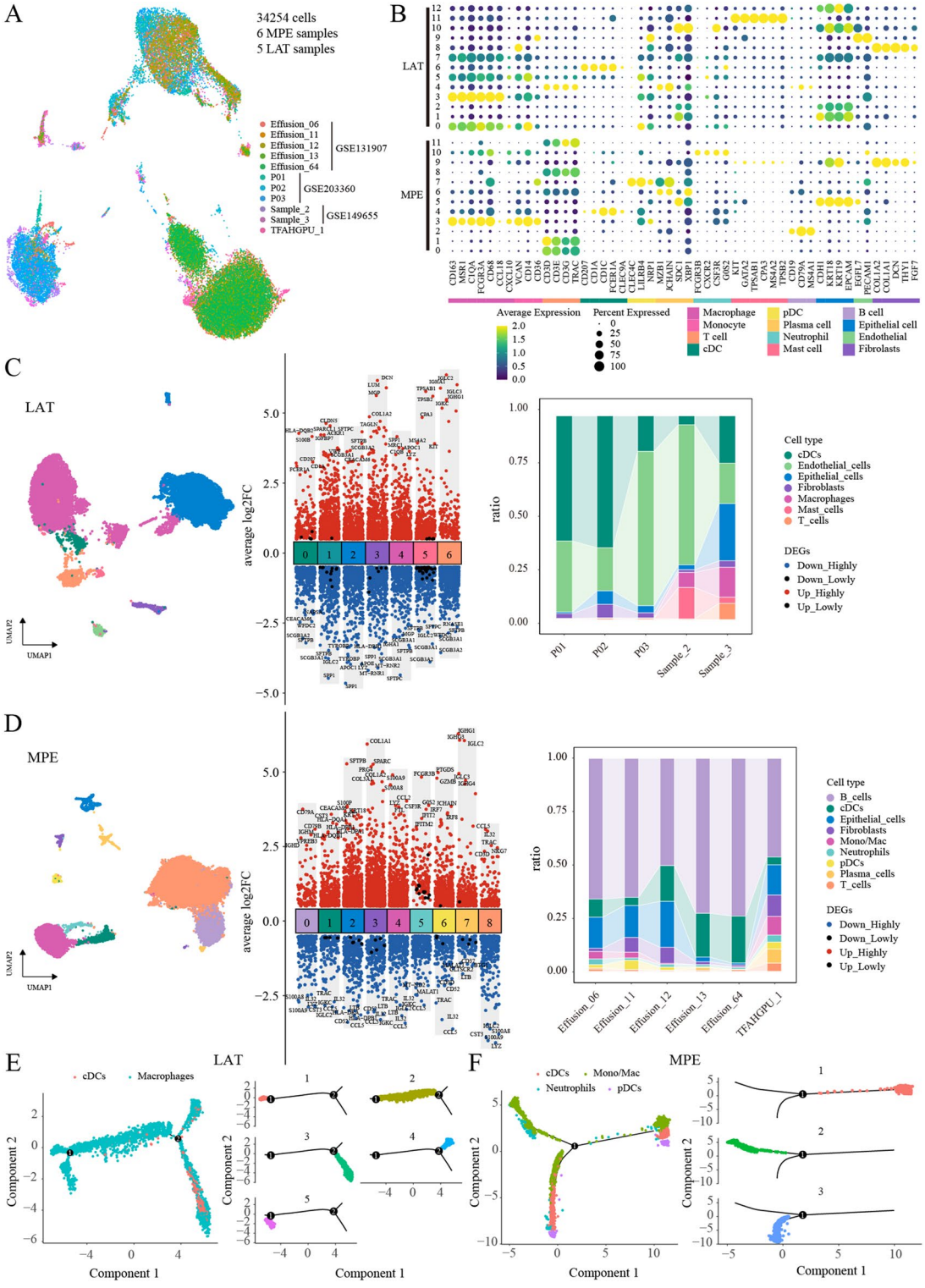

Fig. 2 Single-cell annotation and pseudotime analysis. (**A**) The uniform manifold approximation and projection (UMAP) plot showing the distribution of the 11 samples after removing the batch effects. (**B**) Bubble plot of marker gene expression for cells in lung adenocarcinoma tissue (LAT) and malignant pleural effusion (MPE) samples. (**C**, **D**) UMAP visualization of single-cell annotation results for LAT and MPE samples (left), volcano plots of differentially expressed genes for each cell cluster (middle), and percentage of each cell cluster in each sample (right). Up_Highly: log2FC > 0 and adjusted p value < 0.01; Up_Lowly: log2FC > 0 and adjusted p value >= 0.01, adjusted p value < 0.05; Down_Lowly: log2FC <= 0 and adjusted p value >= 0.01, adjusted p value < 0.05; Down_Highly: log2FC <= 0 and adjusted p value < 0.01. (**E**, **F**) Trajectory analysis of myeloid cells in LAT and MPE samples

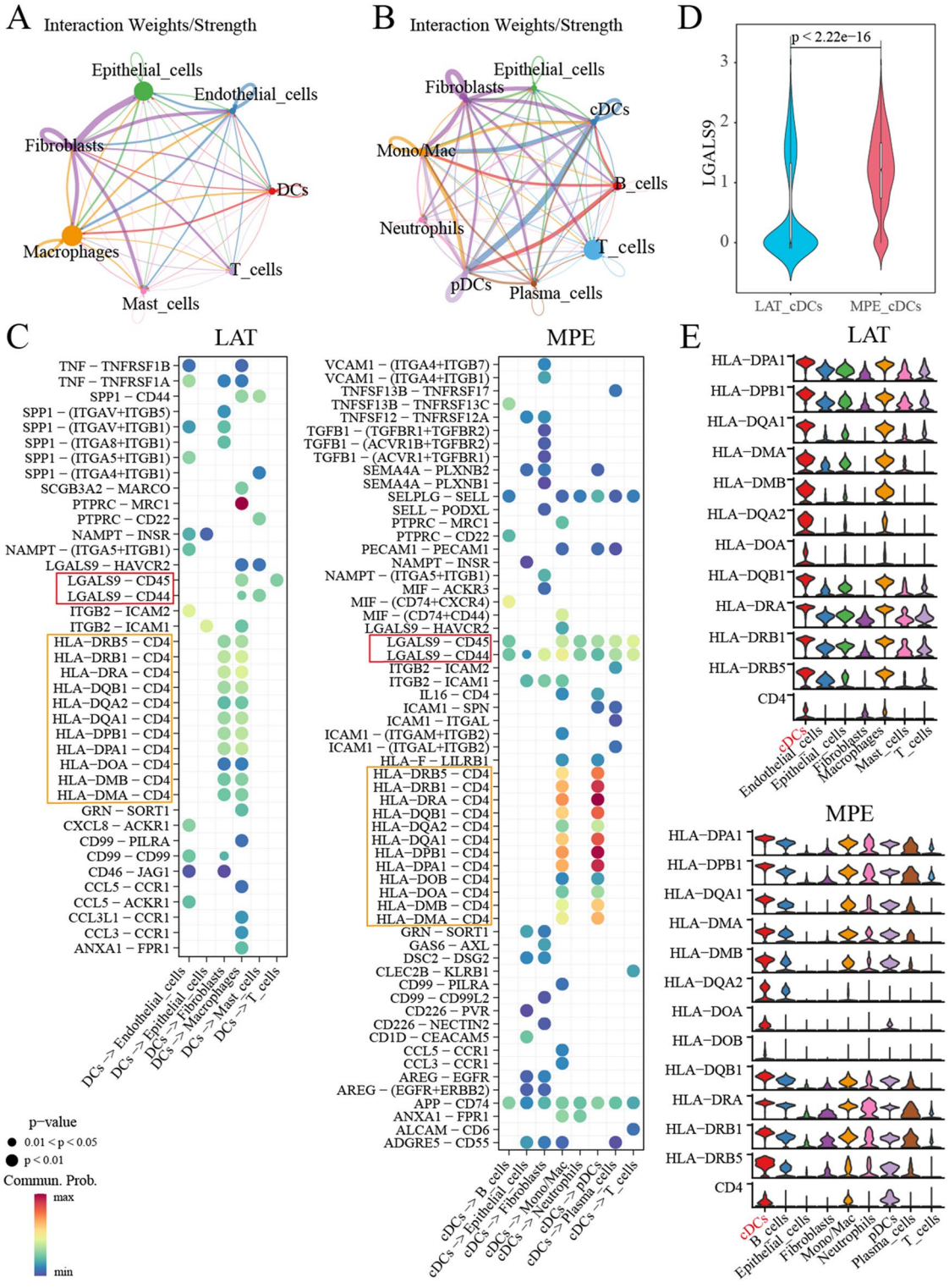

Fig. 3 Analysis of cell-cell communications in single-cell samples. (**A**, **B**) Communication strength between cell clusters in lung adenocarcinoma tissue (LAT) and malignant pleural effusion (MPE) samples. (**C**) Bubble plot of cell-cell communication signals in LAT and MPE samples. (**D**) Violin plot of LGALS9 gene expression in dendritic cells from LAT and MPE samples. (**E**) Violin plot of the expression differences of HLA-D region genes in LAT and MPE samples

to these DC-related genes, leading to the identification of 166 genes that were significantly associated with the prognosis of LUAD patients (Additional file 1: Fig. S1H).

Integration of variables and construction of the prognostic signature

The expression data of 166 DC-related genes were extracted from five LUAD datasets as input data for 101 machine learning models to establish the DCRGS. The C-index, which is widely utilized to assess the accuracy of models, was employed to evaluate the discrimination between the predicted value of the model and the actual value. In this study, by calculating the mean C-index of each model in five LUAD datasets, it was found that the mean C-index of the combination of CoxBoost and RSF models was higher than that of other machine learning model combinations, which indicated that the combination of CoxBoost and RSF models had the best prognostic prediction performance for LUAD patients (Fig. 4A). Twenty feature genes with nonzero coefficients were screened in the CoxBoost algorithm (Fig. 4B), and then the RSF algorithm was used to rank the importance of these genes (Fig. 4C, D). To improve the generalization ability and clinical application convenience of the model, we extracted the top six most important genes, TAP2, PEBP1, PLAUR, STK17B, CXORF21, and MAP3K8. A DCRGS score was generated for each patient by utilizing the expression of the 6 key genes, weighted by their model regression coefficients (Additional file 1: Table S5). Patients were divided into low- and high-DCRGS subgroups using the median DCRGS score of the TCGA-LUAD cohort. Kaplan–Meier survival analysis indicated prominently prolonged OS for the low-DCRGS subgroup in the training and validation cohorts (Fig. 4E-I). The mortality rate increased with increasing DCRGS score in the TCGA and GEO cohorts (Fig. 4J; Additional file 1: Fig. S2), and the expression of TAP2 and PLAUR was positively associated with the DCRGS score, while the expression of PEBP1, STK17B, CXorf21, and MAP3K8 was negatively associated with the DCRGS score (Fig. 4K). We further identified the expression of these six genes within the pseudotime trajectory of myeloid cells. In LAT samples, CXorf21, STK17B, and TAP2 expression were elevated at later stages of the trajectory, whereas PEBP1 and PLAUR were expressed at low levels. However, in MPE samples, PEBP1, STK17B, PLAUR and TAP2 were considerably expressed in the early stage of the trajectory, while MAP3K8 was highly expressed in the later stage (Fig. 4L).

Performance evaluation and application of DCRGS

To experimentally validate the performance of DCRGS, the expression of these six DC-related genes was assessed by RTqPCR in a clinical cohort of 12 patients with LUAD. The DCRGS score of the FAHGPU cohort was calculated and categorized employing a method identical to that of the TCGA-LUAD cohort. The expression of TAP2, PLAUR, and PEBP1 was not substantially different between the two subgroups (Fig. 5A), probably due to the small sample size. However, it was verified that the expression of MAP3K8, STK17B and CXorf21 expression was elevated in the low-DCRGS subgroup. Moreover, time-dependent ROC curves examined the discrimination ability of the DCRGS, and the areas under the curves (AUCs) for 1-year, 2-year, 3-year, 4-year and 5-year survival were 0.705, 0.686, 0.709, 0.734 and 0.686 in TCGA-LUAD; 0.702, 0.671, 0.661, 0.660, and 0.631 in GSE26939; 0.669, 0.604, 0.598, 0.650, and 0.649 in GSE31210; 0.741, 0.673, 0.660, 0.667, and 0.631 in GSE42127; and 0.649, 0.677, 0.663, 0.710 and 0.762 in GSE72094 (Fig. 5B-F). We also compared the predictive capacity of DCRGS with some markers previously identified such as STK11, TP53, KRAS and EGFR mutations, as well as smoking status, sex, age, grade, T stage, N stage, M stage, and AJCC stage. Notably, DCRGS has the best robustness in prognostic assessment of LUAD patients (Fig. 5G). To further illustrate the stability and accuracy of DCRGS for prognostic prediction, we contrasted the predictive performance of DCRGS with that of 85 published prognostic features. Univariate Cox regression analysis was employed to determine the connection between prognostic features and the OS of LUAD patients. DCRGS was substantially associated with OS and a high DCRGS score indicated a profound detrimental effect on the prognosis of patients with LUAD (Additional file 1: Fig. S3A). Furthermore, DCRGS clearly outperformed other prognostic characteristics in all cohorts, according to the comparison of the C index (Additional file 1: Fig. S3B). Most features performed well in their training cohort but poorly in the external cohorts; this bias was caused by the model being overfit.

In the TCGA-LUAD cohort, univariate and multivariate Cox regression analyzes elucidated that DCRGS was an independent prognostic factor (Fig. 6A, B). A nomogram constructed based on DCRGS and clinical features were applied to predict 1-year, 3-year, and 5-year OS rates for the sixth patient of the cohort with AUCs of 0.935, 0.765, and 0.553, respectively (Fig. 6C). Calibration curves, a means of evaluating the performance of the nomogram, indicated significant consistency between the anticipated and actual values of the OS rates for 1 year, 3 year, and 5 years (Fig. 6D). A remarkable discovery was the superior accuracy of DCRGS over other clinical features of the C-index dynamic change profile to predict 5-year OS (Fig. 6E). Decision curve analysis (DCA) demonstrated that the nomogram had an optimal net benefit with a threshold probability between 0.13 and 0.67 (Fig. 6F). Furthermore, we developed an online

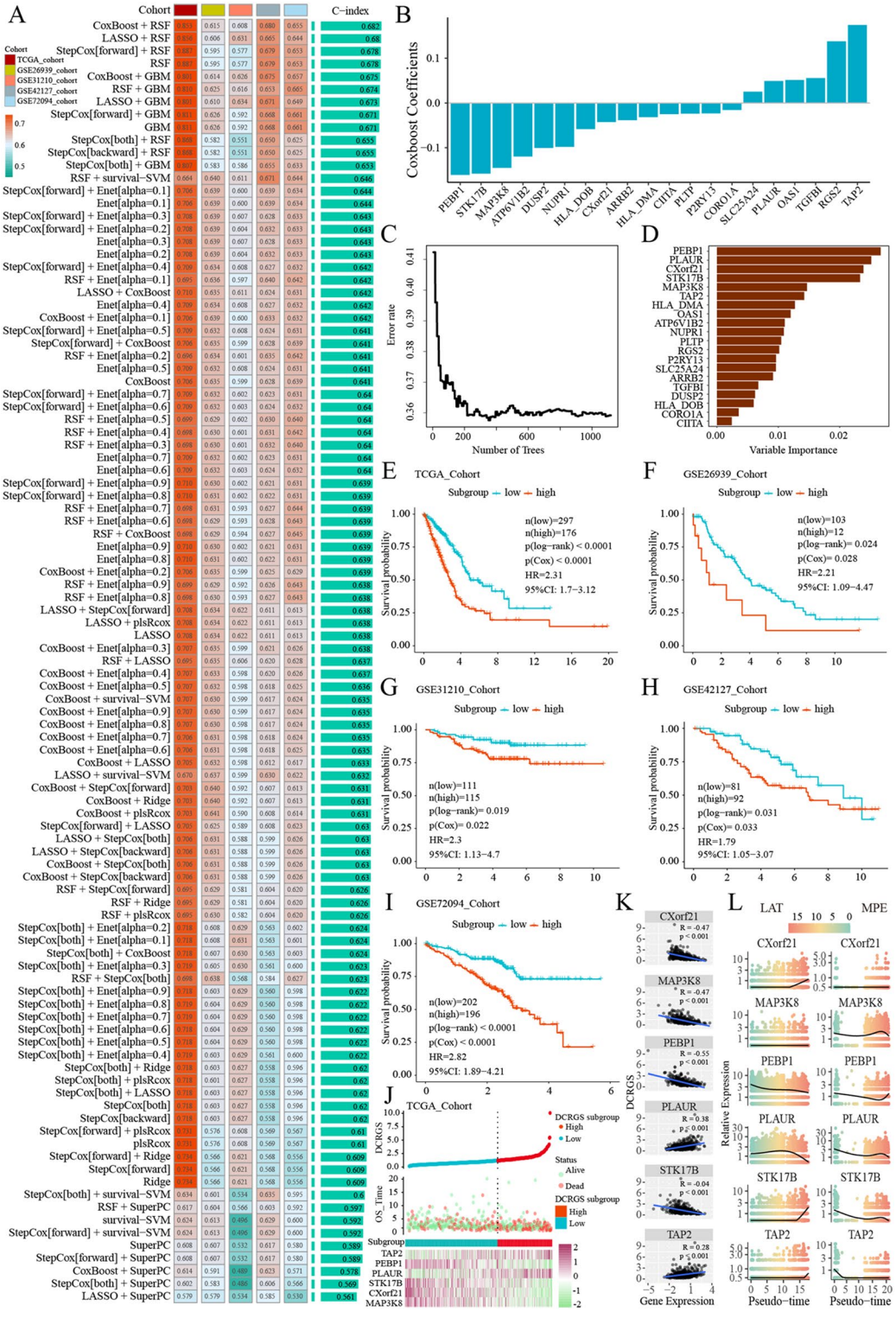

Fig. 4 Machine learning algorithms used to construct DCRGS. (**A**) Heatmap showing the C-index of the 101 machine learning models. (**B**) Visualization of coefficients in Coxboost model. (**C**) Relationship between the number of trees and the error rate in the RSF model. (**D**) Importance ranking of the top 20 genes in the RSF model. (**E-I**) Survival analysis of the low- and high-DCRGS subgroups in five cohorts. (**J**) Risk factor linkage plots demonstrate the overall survival of the TCGA-LUAD cohort and the expression levels of prognostic genes in response to DCRGS. (**K**) Expression levels of the six prognostic genes in the developmental trajectory of myeloid cells. (**L**) Jitter plots showing the expression level of six prognostic genes changing with pseudotime

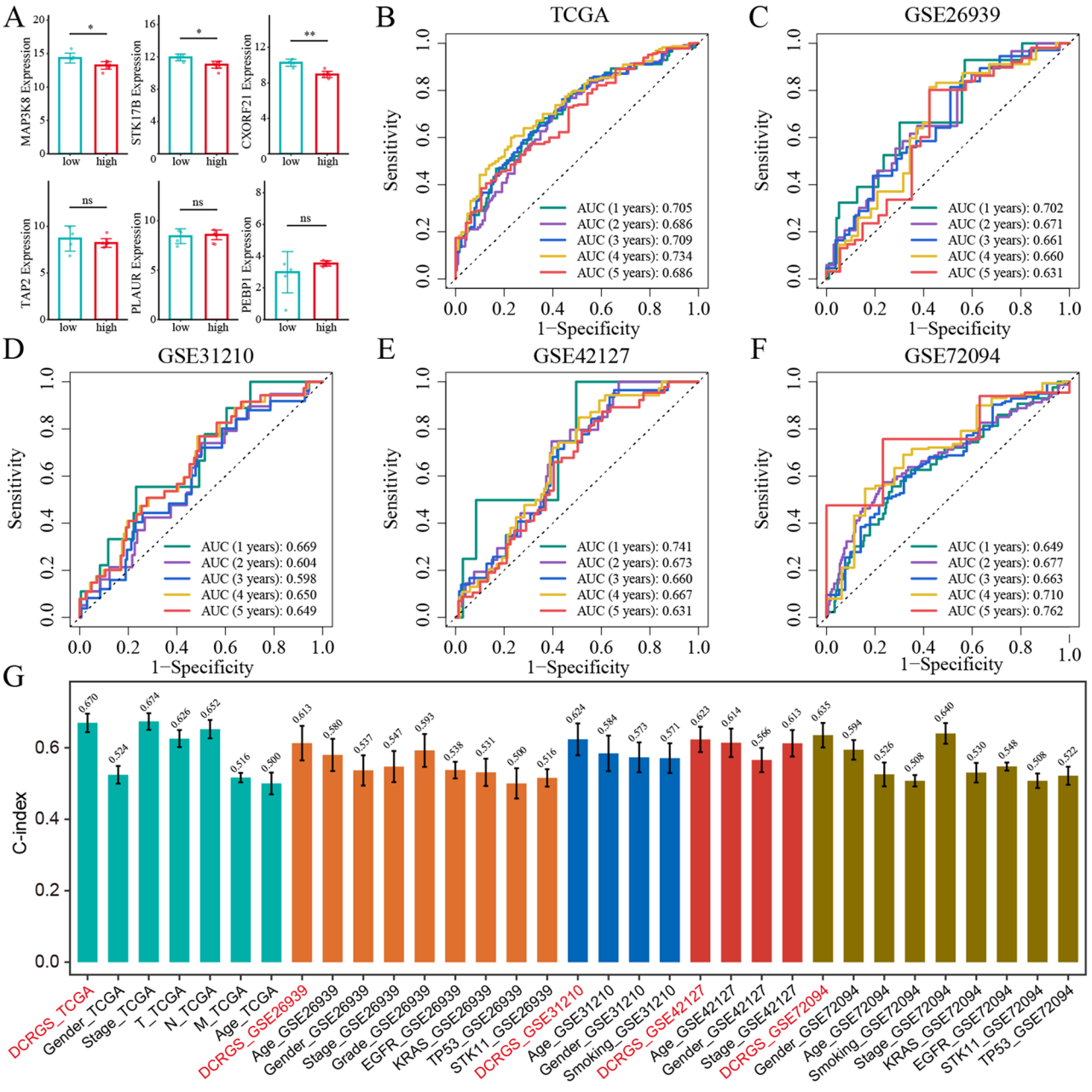

Fig. 5 Performance assessment of DCRGS. (**A**) RT–qPCR was utilized to examine the expression levels of 6 prognostic genes in clinical samples. (**B**-**F**) ROC curves demonstrate the predictive performance of DCRGS in five cohorts. (**G**) C-index bar graph of DCRGS and clinical characteristics of the five cohorts on prognostic assessment. "ns": not significant; "*": $P < 0.05$; "**": $P < 0.01$

application with R software to predict the OS rates of LUAD patients based on their age, sex, T stage, N stage, M stage, and the DCRGS to make the nomogram easier to apply in the clinic (https://kapokshiny.shinyapps.io/DCRGS_LUAD_Nomogram/).

Immune landscape and biological functions of DCRGS

The immune microenvironment plays an important role in tumor development. We used ESTIMATE, xcell, ssGSEA, MCP counts and TIMER2.0 to infer the abundance of infiltrating immune cells between the low- and high-DCRGS subgroups (Fig. 7A). Surprisingly, we discovered that the increase in DCRGS was accompanied by a gradual decline in the immuno-infiltrating abundance. The abundance of DCs, T cells, and macrophages was markedly reduced in the high-DCRGS subgroup, while the abundance of Th2 cells increased. This finding was further confirmed in 4 cohorts from GEO (Additional file 1: Figs. S4, 5).

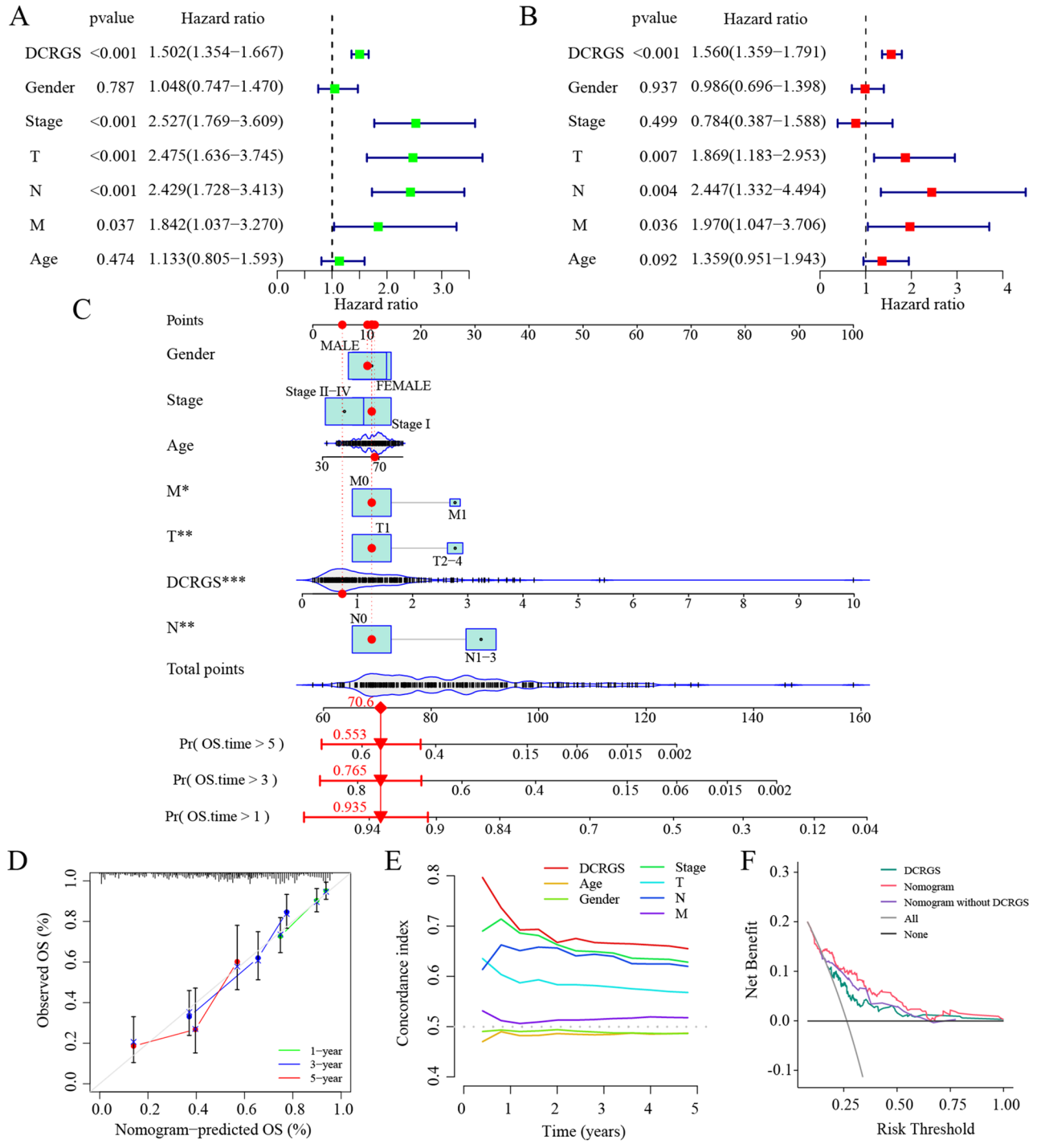

Fig. 6 Construction of the nomogram. (**A**, **B**) Univariate (left) and multivariate (right) Cox regression analysis using DCRGS, gender, stage, T, N and M. (**C**) Nomogram established based on DCRGS and other clinical characteristics. (**D**) Calibration curves assessing the 1-, 3-, and 5-year predictive accuracy of the Nomogram. (**E**) Trend of the C index for 5-year prognostic prediction of clinical features. (**F**) Decision curves of DCRGS and nomogram

To better understand the biological mechanisms underlying the poorer outcomes of patients with high DCRGS, 15 tumor-associated gene sets were utilized for ssGSEA, which revealed that DNA damage repair pathways (p53 signaling, mismatch repair, homologous recombination) and oncogenic pathways (Wnt signaling, cell cycle) were significantly upregulated in the high-DCRGS subgroup, whereas immune-associated pathways (T-cell and B-cell receptor signaling, Fcγ R-mediated phagocytosis) were considerably enriched in the low-DCRGS subgroup

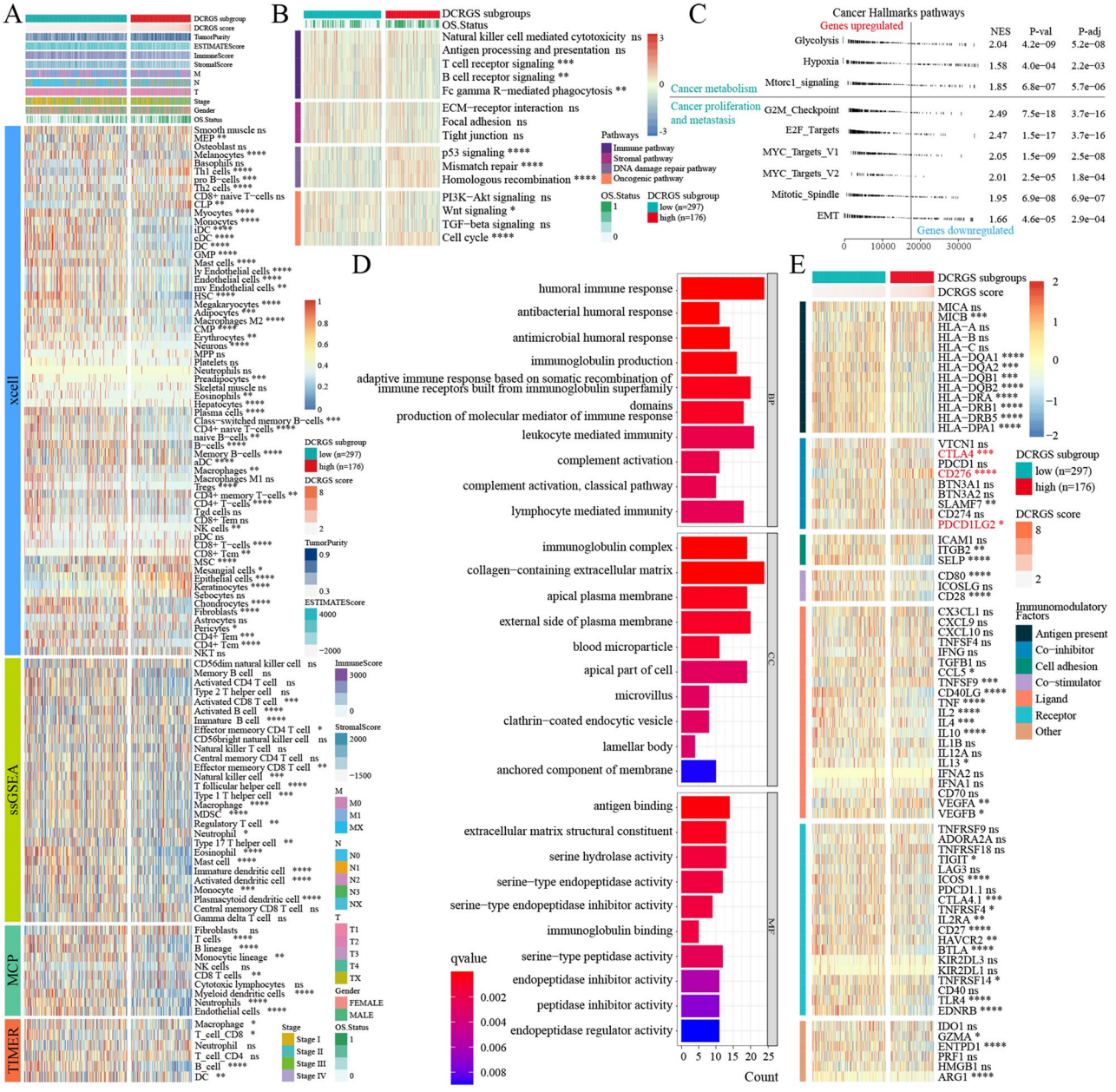

Fig. 7 Biological characterization of the low and high DCRGS subgroups. (**A**) Five algorithms inferred the infiltration abundance of immune cells. (**B**) Heatmap of the expression levels of 15 pathways in the low- and high-DCRGS subgroups. (**C**) Enrichment levels of Hallmarks pathways associated with tumor metabolism, proliferation and metastasis in both subgroups. (**D**) GO enrichment analysis of Molecular Function (MF), Biological Process (BP) and Cell Component (CC) in both subgroups. (**E**) Heatmap of the expression levels of immunomodulatory factors in the two subgroups. "ns": not significant; "*": $P < 0.05$; "**": $P < 0.01$; "***": $P < 0.001$; "****": $P < 0.0001$

(Fig. 7B). Furthermore, the high-DCRGS subgroup showed a strong activation of the tumor metabolism, proliferation, and metastasis-related pathways obtained from the hallmark gene sets (Fig. 7C). To learn more about the molecular mechanisms by which DCRGS influenced the prognosis, GO enrichment analysis was performed that included the biological process (BP), cellular component (CC) and molecular function (MF). The top 10 GO terms in both subgroups were mostly linked to biological processes involving humoral immunity (Fig. 7D). Immunomodulatory factor expression levels were also strongly correlated with DCRGS. In the high-DCRGS subgroup, the expression levels of HLA-D-region genes, cell adhesion molecules, costimulatory molecules, and ligand- and receptor-related molecules were all noticeably decreased, but the expression of the coinhibitory molecules CD276 and PDCD1LG2 was elevated (Fig. 7E). Surprisingly, the low-DCRGS subgroup had dramatically enhanced

CTLA-4 expression, indicating that these patients would benefit more from anti-CTLA-4 ICI treatment. Furthermore, the GEO cohort supported these results (Additional file 1: Fig. S6).

SNV data were employed to examine the distribution of somatic mutations in both subgroups, and 20 highly mutated genes associated with DCRGS were identified (Fig. 8A). Fisher's test identified 18 differentially mutated genes ($p < 0.01$) (Fig. 8B) with higher mutation rates in the high-DCRGS subgroup, including TP53 (63%), RP1L1 (25%), PLPPR4 (17%) and PCSK5 (9%) (Fig. 8C). We found that TMB was positively correlated with the DCRGS score and that TMB was significantly higher in the high DCRGS subgroup (Fig. 8D). All differentially mutated genes were subjected to GSEA to further elucidate biological characteristics (Fig. 8E). The results

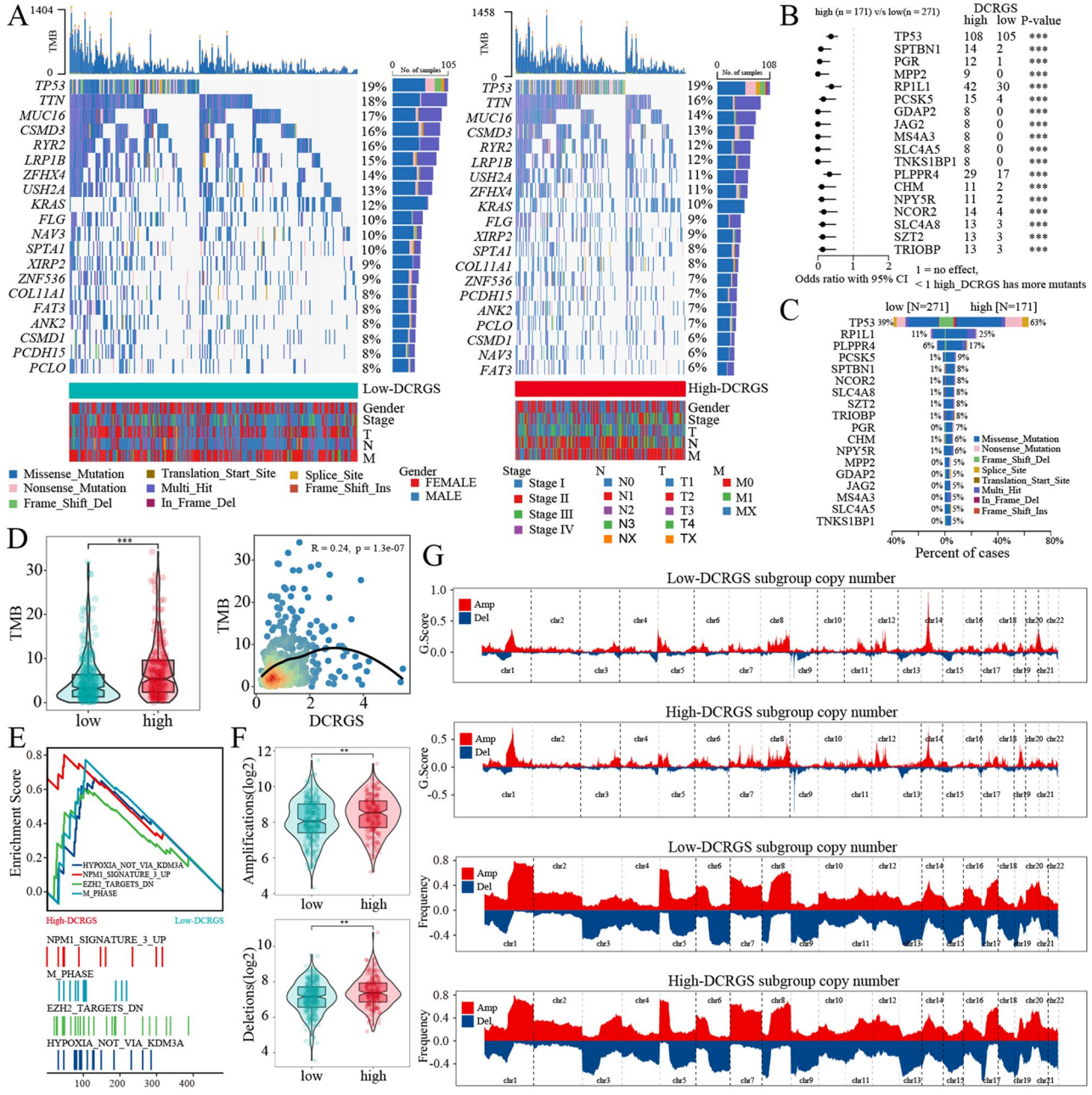

Fig. 8 Analysis of somatic mutation and copy number variation (CNV). (**A**) Waterfall plots of the top 20 mutated genes in the low- and high-DCRGS subgroups. (**B**, **C**) Forest and bar graphs of differentially mutated genes in both subgroups. (**D**) Variations in the mutation loads of the two subgroups (left) and the connection between the mutation loads and DCRGS (right); denser points are shown by redder colors in the plot. (**E**) GSEA of differentially mutated genes. (**F**) Violin plots of amplification and deletion mutations between the two subgroups. (**G**) CNV plot demonstrating the gistic score and the mutation frequency distribution in both subgroups

revealed that the NPM1 signature, M phase, EZH2 targets, and HYPOXIA not via KDM3A pathways were significantly enriched. These pathways facilitate immunological remodeling, inhibition of antigen presentation, suppression of immune cell activity, and cell division, which promote immune escape and tumor growth [30–32]. Given that CNVs might cause genome variability, the correlation between DCRGS and CNV was further examined. A striking increase in chromosomal amplifications and deletions was observed in the high-DCRGS subgroup (Fig. 8F). To further clarify the CNVs occurring on each chromosome, Fig. 8G depicts the distribution of CNV in both subgroups.

Immunotherapy sensitivity prediction and drug screening

TCIA was used to collect medication data for the TCGA-LUAD cohort to better understand the connection between DCRGS and drug sensitivity. Immunotherapy response analysis revealed an elevated sensitivity to anti-CTLA4 monotherapy in the low-DCRGS subgroup, but there were no discernible differences in either subgroup for the combination of anti-PD-1 and anti-CTLA4 or anti-PD-1 monotherapy (Fig. 9A). Interestingly, the high-DCRGS subgroup responded better to paclitaxel (Fig. 9B). Next, we analyzed the relationship between the prognostic signature and the traditional therapeutic targets of LUAD (Fig. 9C). To further identify potential target drugs, we compared the half-maximum inhibitory concentration of each drug (IC50) in both subgroups and identified those with notable differences ($p < 0.001$). Figure 9D explains the relationship between drug sensitivity and DCRGS, with the eight most sensitive drugs highlighted. Figure 9E demonstrates the correlation between these drugs and DCRGS. We found that the IC50s of foretinib, MK-1775, dasatinib, BI-2536, PD0325901 and AZD7762 were lower in the high-DCRGS subgroup, while the IC50s of BMS-754,807 and AZD8055 were higher.

Discussion

In the present study, scRNA-seq and bulk-seq data were utilized to elucidate the function of DCs within the tumor microenvironment. A DCRGS was developed with machine learning algorithms to predict the prognosis and response to immunotherapy in patients with LUAD. A significant upregulation of the LGALS9-CD44/CD45 ligand-receptor pair signaling was detected in DCs within MPE, which is closely correlated with an unfavorable prognosis in LUAD patients. Furthermore, by implementing machine learning algorithms, we have identified six crucial DC-related genes. Based on these genes, we subsequently developed a prognosis prediction model, DCRGS, capable of predicting the prognosis of LUAD patients. The stability and robustness of our DCRGS were validated both internally and externally, exhibiting excellent performance. Compared to 85 previously published prognostic characteristics for LUAD patients, our DCRGS demonstrated a superior generalizability. Furthermore, our analysis elucidated that the DCRGS possesses the capability to evaluate the inhibitory status of TME, ascertain the activity of tumor-associated pathways, distinguish between TMB and CNV load levels, and recognize potential therapeutic agents. Consequently, derived from these findings, our study holds significant implications for advancing the field of prognosis prediction and personalized medicine in the context of LUAD.

DCs, one type of specialized APCs, have substantially higher antigen-presenting activity than macrophages and B cells and build an important bridge between adaptive and innate immunity [33]. Previous studies have demonstrated that a significant decrease in the activity of cytotoxic T lymphocyte-mediated antitumor activity is linked to LGALS9 overexpression in DCs, which has been shown to limit the functions of antigen recognition, processing, and presentation in DCs [34, 35]. The possibility of communication between cDCs and T cells via the LGALS9-CD44/CD45 ligand−receptor pair was significantly higher in the MPE samples of this study than in the LAT samples, which strongly explains the poor prognosis of LUAD patients who suffered from MPE. Furthermore, by applying CellChat, we revealed that HLA-D region gene expression was significantly elevated in DCs from LAT and MPE samples and engaged in macrophage-DC interactions. The antigen presentation function of intratumor HLA-II is dominated primarily by dedicated APCs [36]. DCs interact with CD169 + macrophages, which are responsible for capturing antigens, and then present the picked-up antigens to T cells, thereby activating them [37]. These results indicate that DCs play an important role in TME.

Through the combination of CoxBoost and RSF, a set of six pivotal genes (CXorf21, MAP3K8, PEBP1, PLAUR, STK17B, and TAP2) have been identified as significant prognostic indicators for patients with LUAD. CXorf21, an immune-related gene, was renamed Toll-like receptor adapter interacting with SLC15A4 in the lysosome (TASL), which is expressed in B cells, CD33 + monocytes, dendritic cells and macrophages [38, 39]. MAP3K8 is a LUAD transforming gene whose aberrant transcriptional regulation, gene amplification, and mutations are implicated in the development of a variety of cancers, including thymoma, lymphoma, breast cancer, nasopharyngeal carcinoma and lung cancer [40]. According to some research, MAP3K8 can be utilized as a prognostic marker for LUAD patients [41]. The value of MAP3K8 in predicting the prognosis of LUAD patients was also verified by this study. PEBP1, known as RAF kinase

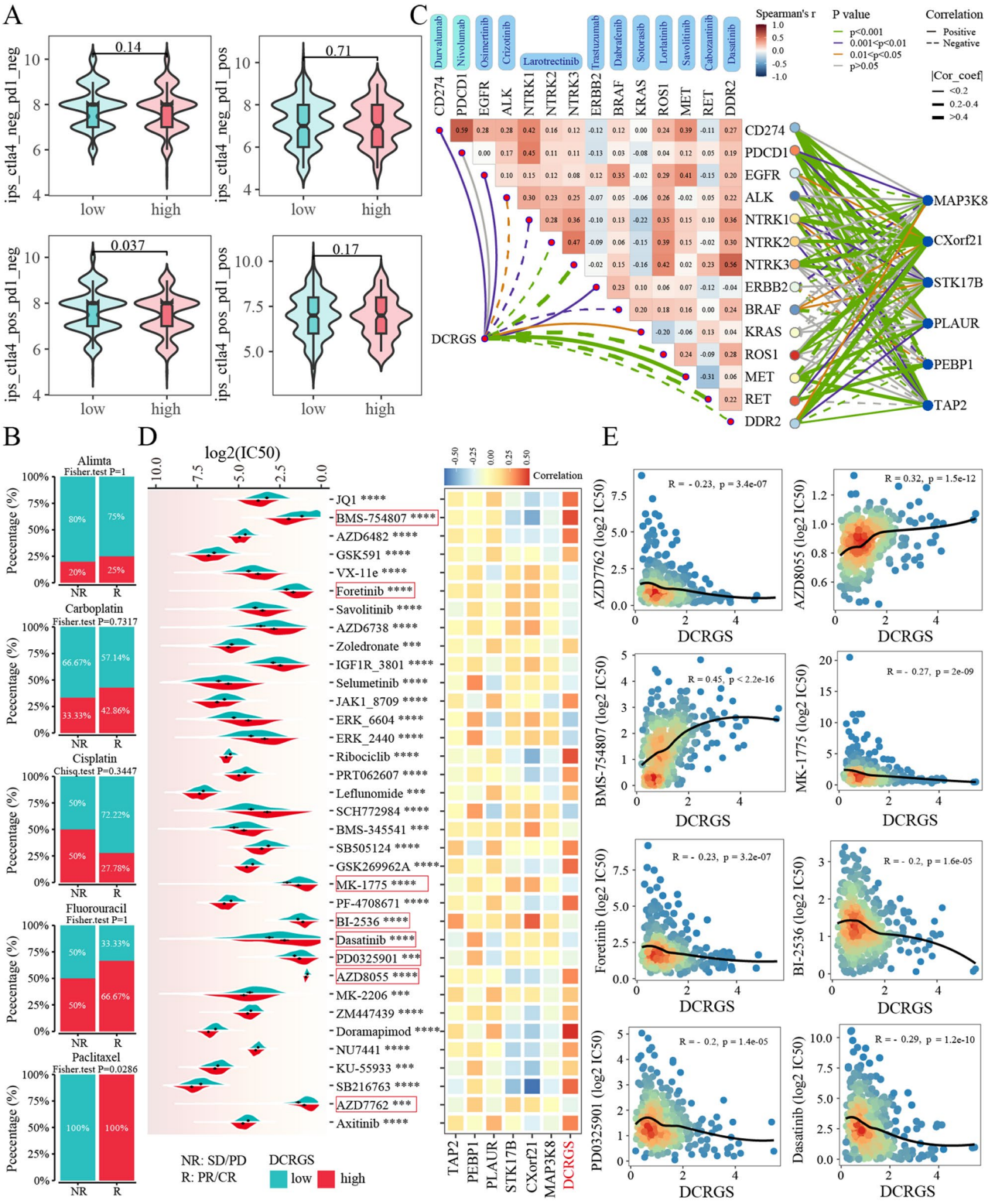

Fig. 9 Analysis of immunotherapy sensitivity and drug screening. (**A**) Comparison of sensitivity to anti-CTLA-4 and anti-PD-1 inhibitors in low- and high-DCRGS subgroups. (**B**) Differences in response to chemotherapeutic agents in both subgroups. NR: no response; SD: stable disease; PD: progressive disease; R: response; PR: partial response; CR: complete response. (**C**) Relationship of the commonly used target genes for LUAD with prognostic genes and DCRGS. (**D**) Prediction of the drug IC50 in the two subgroups. (**E**) Correlation between IC50 and DCRGS. The higher the density of dots, the redder the color. "***": $P < 0.001$; "****": $P < 0.0001$

inhibitory protein (RKIP), is recognized as a tumor suppressor [42, 43], antagonizes the activity of the B-RAF isoform kinase [44] and inhibits a variety of tumor cell biological processes, including tumor cell growth and invasion [45–47], epithelial to mesenchymal transition (EMT) [48] and metastasis [49]. By inhibiting the HMGA2 signaling pathway, PEBP1 expressed in the TME suppresses macrophage chemokine production, which in turn reduces the infiltration abundance and metastatic potential of TAMs [50]. Furthermore, PEBP1, a novel effector of the apoptosis-inducing signaling pathway, was found to be a prognostic marker for LUAD patients [51–53]. Low expression of PEBP1 is related to tumor metastasis and the poor prognosis. By activating plasma fibrinogen and degrading the extracellular matrix, PLAUR contributes to biological activities such as cell migration, adhesion, and proliferation [54, 55]. Inhibiting PLAUR expression can also suppress the metastasis of lung cancer to lymph nodes [56]. The progression of acute myeloid leukemia [57], chronic lymphocytic leukemia [58] and colorectal cancer [59] is suppressed by upregulation of STK17B, a member of the death-associated protein kinase (DAPK) family [60]. However, the function of STK17B in other cancers is highly contested. Hepatocellular carcinoma (HCC) has been shown to have a greatly increased STK17B expression, and suppression of STK17B expression significantly inhibits HCC growth and metastasis [61]. The tumorigenic ability of cancer cells is inhibited when STK17B expression is reduced in breast cancer [62]. This study discovered that patients in the low-DCRGS subgroup had a higher STK17B expression and noticeably longer OS than those in the high-DCRGS subgroup, illustrating that low STK17B expression is associated with a poor prognosis in LUAD patients and that reduced STK17B expression can be identified as a biomarker of poor prognosis in LUAD patients. TAP2 is one of the crucial genes in the MHC class I antigen presentation pathway. The protein encoded by TAP2 serves the processing of endogenous antigens. An important mechanism of tumor immune escape is the increased mutation frequency of TAP2 with altered DNA damage response and repair (DDR), which in turn causes the antigen presentation efficacy to decline [63]. We revealed that patients with high expression of TAP2 exhibited a worse prognosis than those with low expression. MAP3K8, STK17B, and CXorf21 expression trends in the low- and high-DCRGS subgroups were initially confirmed by RTqPCR to be compatible with the findings of the bioinformatics analysis. The main source of bias may be the low number of LAT samples, given that the expression levels of TAP2, PLAUR, and PEBP1 in both subgroups were not clearly distinguished.

Six key genes were used to create a novel prognostic signature for prognostic signature for LUAD called DCRGS, and four distinct GEO cohorts were used to evaluate its robustness and stability for the prognostic prediction of LUAD patients. Additionally, clinical features, such as STK11, TP53, KRAS, and EGFR mutations, smoking status, sex, age, stage, T stage, N stage, and M stage, had a prominent impact on the prognosis of LUAD patients; however, it is noteworthy that the DCRGS operated independently of these clinical features. The findings of the C-index assessment showed that DCRGS performed better compared to these clinical features for prognostic prediction. Additionally, based on the DCRGS, we built an online interactive tool for predicting LUAD patient prognosis in a more practical and understandable manner.

To better understand the prognostic value of DCRGS, we separated the TCGA-LUAD cohort into low- and high-DCRGS subgroups and examined the immune cell infiltration in each subgroup. The abundance of infiltrating DCs, macrophages, and CD8 + T cells in the high-DCRGS subgroup was generally lower than that in the low-DCRGS subgroup, according to five algorithmic extrapolations. Both DCs and macrophages are APCs, but DCs with higher antigen presentation efficiency benefit from being professional APCs. Macrophages phagocytose and process tumor antigens, collaborating with DCs to deliver antigens to T cells. Then, effector T cells exert antitumor effects, with CD8 + T cells being a crucial component of this process. In this study, enrichment analysis of tumor-associated pathways revealed that the high-DCRGS subgroup was considerably enriched for pathways involved in DNA damage repair, oncogenesis and tumor metabolism, proliferation and metastasis. GO enrichment analysis showed that differentially expressed genes were enriched for immune-related terms. Furthermore, in the TME of the high-DCRGS subgroup, the expression of coinhibitory molecules was substantially increased, whereas that of antigen presentation, cell adhesion factors, and costimulatory molecules was markedly reduced. However, these elements are responsible for the development of immune rejection or immune desert [64, 65]. We also discovered a correlation between DCRGS and both the somatic mutation burden and the CNV load. Differentially mutated genes in the high-DCRGS subgroup were markedly enriched in tumor cell division, growth, invasion, and immune escape pathways. These findings provide convincing evidence for the biological factors driving the poor prognosis of patients with high DCRGS, as well as the value of DCRGS in predicting patient prognosis.

LUAD patients with driver mutations show varying degrees of benefit from targeted therapy, but most develop resistance soon after administration [66, 67], at which point patients face the dilemma of being drug-free. Immunotherapy has become a new breakthrough

in tumor treatment; however, only a small proportion of patients with LUAD receive long-term clinical benefits, and the majority of patients experience disease progression due to primary or acquired resistance during or after ICI discontinuation [68]. Drug sensitivity analysis of TCGA-LUAD patients based on DCRGS revealed that patients with low DCRGS showed a favorable response to anti-CTLA4 therapy, which may be associated with higher levels of immune cell infiltration abundance and CTLA4 expression in the low-DCRGS subgroup. Several studies have demonstrated that patients with massive infiltration of immune cells, especially those with very high levels of T cells, responded better to ICI treatment [69, 70]. Furthermore, we found that the high-DCRGS subgroup has a favorable response to chemotherapy besides targeted therapy. Notably, there is a subtle correlation between DCRGS and the oncogenic drivers of LUAD. DCRGS shows the positive correlations with expression of KRAS and ERBB2. Studies have illustrated that overexpression of ERBB2 and KRAS in LUAD is linked to tumor invasion and adverse prognosis [71, 72].

The DCRGS signature, constructed by machine learning algorithms, can be executed to predict the prognosis of LUAD patients and potential drugs by detecting the expression of six key genes. Although DCRGS shows prominent convenience and accuracy with high clinical translational value, this study still has limitations. First, more research is needed to understand how LGALS9, which is highly expressed in MPE samples, facilitates the poor prognosis of LUAD patients. Second, before applying the findings of big data analysis to the clinic, an adequate number of clinical samples must be used to further confirm them. Finally, the efficacy and safety of several medications still need to be examined in clinical trials, although the treatments suggested in this study offer more options for patients with LUAD, particularly those who have developed drug resistance.

Conclusion

In summary, we identify DC-related genes using scRNA-seq and bulk-seq and develop a machine learning-based prognostic signature consisting of the top six most important DC-related genes, i.e. TAP2, PEBP1, PLAUR, STK17B, CXORF21, and MAP3K8. The model we developed, DCRGS, possesses an exceptional prognosis predictive capacity for patients with LUAD, as validated by numerous independent datasets. DCRGS can also be used to identify the immune status of TME and assist in screening drugs for LUAD patients. Interestingly, we unexpectedly discovered that the signaling of the LGALS9-CD44 / CD45 receptor pair was dramatically activated in MPE samples, which is correlated with a poor prognosis. We believe that the results of this study offer a wealth of valuable information on clinical prognostic assessment, quantitative risk management, and personalized clinical treatment of LUAD patients.

Abbreviations

LUAD	Lung adenocarcinoma
DC	Dendritic cell
scRNA-seq	Single-cell RNA sequencing
DCRGS	DCs-related gene signature
ROC	Receiver operating characteristic
NSCLC	Non-small cell lung cancer
TMB	Tumor mutational burden
TME	Tumor microenvironment
APC	Antigen presenting cell
MHC-I	Major histocompatibility complex class I
MPE	Malignant pleural effusion
bulk-seq	Bulk RNA sequencing
LAT	Lung adenocarcinoma tissue
OS	Overall survival
TCGA	The Cancer Genome Atlas
GEO	Gene Expression Omnibus
RTqPCR	Real-time quantitative PCR
GSEA	Gene set enrichment analysis
MSigDB	Molecular Signatures Database
TCIA	The Cancer Immunome Atlas
UMIs	Unique molecular identifiers
WGCNA	Weighted gene coexpression network analysis
TOM	Topological overlap matrix
LASSO	Least absolute shrinkage and selection operator
RSF	Random survival forest
Enet	Elastic network
plsRcox	Partial least squares regression for Cox
SuperPC	Supervised principal components
GBM	Gradient boosting machine
Survival-SVM	Survival support vector machine
StepCox	Stepwise Cox regression
LOOCV	Leave-one-out cross-validation
C-index	Concordance index
ssGSEA	Single-sample gene set enrichment analysis
GO	Gene Ontology
SNV	Single-nucleotide variant
CNV	Copy number variant
IC50	Half maximal inhibitory concentration
AUC	Areas under the curves
DCA	Decision curve analysis
RKIP	RAF kinase inhibitory protein
EMT	Epithelial to mesenchymal transition
DAPK	Death-associated protein kinase
HCC	Hepatocellular carcinoma
DDR	Damage response and repair
TKI	Tyrosine kinase inhibitor
BP	Biological process
CC	Cellular component
MF	Molecular function

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12935-025-03642-z.

Supplementary Material 1: Table S1. List of gene sets including 14 biological processes. Table S2. Grouping information from five cohorts. Table S3. 85 LUAD prognostic features from PubMed. Table S4. Primer sequences of six key genes. Table S5. Coefficients of signature genes in regression models. Fig. S1. Analysis of the weighted gene coexpression network for the identification of dendritic cell-associated genes. (A) Survival analysis of groups based on the best cutoff value for dendritic cells (DCs) infiltration abundance. (B) Soft threshold versus scale-free fit index (left) and mean connectivity (right). (C) Clipping of clustered modules below 0.3. (D) Dynamic modules clipping and merging of modules. (E) Heatmap of coexpression between module genes. (F) Correlation between immune

infiltration abundance of DCs and module gene expression. (G) Intersection of DCs-related genes filtered using scRNA-seq and bulk-seq data. The intersection of two gene sets is tested with a hypergeometric distribution. (H) Volcano plot of the relationship between Z-score and negative p-value logarithm of DCs-related genes associated with prognosis. Fig. S2. Risk factor linkage plots of four validation cohorts. Fig. S3. Comparison of prognostic signatures for LUAD. (A) P-values and hazard ratios (HR) of the univariate Cox regression analysis of DCRGS and 85 published signatures in five cohorts. (B) C-index for assessing the predictive performance of DCRGS and 85 published signatures in five cohorts. Fig. S4. Immune infiltration analysis of the GSE26939 and GSE31210 cohorts. "ns": not significant; "*": $P < 0.05$; "**": $P < 0.01$; "***": $P < 0.001$; "****": $P < 0.0001$. Fig. S5. Immune infiltration analysis of the GSE42127 and GSE72094 cohorts. "ns": not significant; "*": $P < 0.05$; "**": $P < 0.01$; "***": $P < 0.001$; "****": $P < 0.0001$. Fig. S6. Immunomodulatory factor expression profiles in the four cohorts. "ns": not significant; "*": $P < 0.05$; "**": $P < 0.01$; "***": $P < 0.001$; "****": $P < 0.0001$

Acknowledgements
Not applicable.

Author contributions
Z.S, J.B., S.C. and Y.L. contributed to the conception of the study. Z.S., M.H., J.B., X.H., M.S., X.C. and S.C. contributed to collection and organization of data; Z.S., M.H., S.C. and Y.L. contributed to data analysis; Z.S., M.H., J.B. and X.H. contributed to data interpretation; Z.S. completed manuscript writing. J.B., S.C., and Y.L. contributed to the substantive revision of the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding
This study was supported by grants from the Guangdong Provincial Department of Science and Technology (2020A0505100058), Natural Science Foundation of China (32271225), Guangdong Provincial Medical Products Administration (2022TDZ19), Guangzhou Basic and Applied Basic Research Foundation (2024A04J4811), and Hospital Fund of the First Affiliated Hospital of Guangdong Pharmaceutical University (ZXYXTQJ202304).

Data availability
The corresponding author can provide any non-open-source data and R code used in this study upon justifiable request. The public data used in this study were obtained from TCGA (https://portal.gdc.cancer.gov), MSigDB (https://www.gsea-msigdb.org/), TCIA (https://tcia.at), and GEO (https://www.ncbi.nlm.nih.gov/geo/).

Declarations

Ethics approval and consent to participate
The study was carried out in accordance with the Declaration of Helsinki and was approved by the Human Ethics Committee of Guangdong Pharmaceutical University's First Affiliated Hospital (approval No. 69, 2022).

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of Immuno-Oncology, The First Affiliated Hospital of Guangdong Pharmaceutical University, Guangzhou 510080, China
[2]Guangdong Provincial Engineering Research Center for Esophageal Cancer Precision Therapy, The First Affiliated Hospital of Guangdong Pharmaceutical University, Guangzhou 510080, China
[3]Department of Internal Medicine, The First Affiliated Hospital of Anhui University of Chinese Medicine, Hefei 230000, China
[4]Key Laboratory of Cancer Immunotherapy of Guangdong Higher Education Institutes, Guangzhou 510080, China
[5]Research & Development Division, Guangzhou Anjie Biomedical Technology Co., Ltd., Guangzhou 510535, China
[6]Key Laboratory of Monitoring Adverse Reactions Associated with CAR-T Cell Therapy, Guangzhou 510080, China

Received: 25 July 2024 / Accepted: 7 January 2025
Published online: 14 January 2025

References

1. Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A, Bray F. Global Cancer statistics 2020: GLOBOCAN estimates of incidence and Mortality Worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2021;71(3):209–49.
2. Gibney GT, Weiner LM, Atkins MB. Predictive biomarkers for checkpoint inhibitor-based immunotherapy. Lancet Oncol. 2016;17(12):e542–51.
3. Sharma P, Hu-Lieskovan S, Wargo JA, Ribas A. Primary, adaptive, and Acquired Resistance to Cancer Immunotherapy. Cell. 2017;168(4):707–23.
4. Zhang J, Fan B, Cao G, Huang W, Jia F, Nie G, Wang H. Direct presentation of Tumor-Associated antigens to induce adaptive immunity by personalized dendritic cell-mimicking nanovaccines. Adv Mater. 2022;34(47):e2205950.
5. Fu C, Jiang A. Dendritic cells and CD8 T cell immunity in Tumor Microenvironment. Front Immunol. 2018;9:3059.
6. Vonderheide RH. The Immune Revolution: a case for Priming, not checkpoint. Cancer Cell. 2018;33(4):563–9.
7. Goc J, Germain C, Vo-Bourgais TK, Lupo A, Klein C, Knockaert S, de Chaisemartin L, Ouakrim H, Becht E, Alifano M, et al. Dendritic cells in tumor-associated tertiary lymphoid structures signal a Th1 cytotoxic immune contexture and license the positive prognostic value of infiltrating CD8 + T cells. Cancer Res. 2014;74(3):705–15.
8. Mansuet-Lupo A, Alifano M, Pécuchet N, Biton J, Becht E, Goc J, Germain C, Ouakrim H, Régnard JF, Cremer I, et al. Intratumoral Immune Cell densities are Associated with Lung Adenocarcinoma Gene alterations. Am J Respir Crit Care Med. 2016;194(11):1403–12.
9. Barry KC, Hsu J, Broz ML, Cueto FJ, Binnewies M, Combes AJ, Nelson AE, Loo K, Kumar R, Rosenblum MD, et al. A natural killer-dendritic cell axis defines checkpoint therapy-responsive tumor microenvironments. Nat Med. 2018;24(8):1178–91.
10. Hanabuchi S, Watanabe N, Wang YH, Wang YH, Ito T, Shaw J, Cao W, Qin FX, Liu YJ. Human plasmacytoid predendritic cells activate NK cells through glucocorticoid-induced tumor necrosis factor receptor-ligand (GITRL). Blood. 2006;107(9):3617–23.
11. Salio M, Cella M, Vermi W, Facchetti F, Palmowski MJ, Smith CL, Shepherd D, Colonna M, Cerundolo V. Plasmacytoid dendritic cells prime IFN-gamma-secreting melanoma-specific CD8 lymphocytes and are found in primary melanoma lesions. Eur J Immunol. 2003;33(4):1052–62.
12. Zhang XC, Wang J, Shao GG, Wang Q, Qu X, Wang B, Moy C, Fan Y, Albertyn Z, Huang X, et al. Comprehensive genomic and immunological characterization of Chinese non-small cell lung cancer patients. Nat Commun. 2019;10(1):1772.
13. De Zuani M, Xue H, Park JS, Dentro SC, Seferbekova Z, Tessier J, Curras-Alonso S, Hadjipanayis A, Athanasiadis EI, Gerstung M, et al. Single-cell and spatial transcriptomics analysis of non-small cell lung cancer. Nat Commun. 2024;15(1):4388.
14. Mahmood K, Jampani P, Clarke JM, Wolf S, Wang X, Wahidi MM, Giovacchini CX, Dorry M, Shofer SL, Shier J, et al. High yield of Pleural Cell-Free DNA for diagnosis of oncogenic mutations in Lung Adenocarcinoma. Chest. 2023;164(1):252–61.
15. Tong L, Ding N, Tong X, Li J, Zhang Y, Wang X, Xu X, Ye M, Li C, Wu X, et al. Tumor-derived DNA from pleural effusion supernatant as a promising alternative to tumor tissue in genomic profiling of advanced lung cancer. Theranostics. 2019;9(19):5532–41.
16. Collins FS, Morgan M, Patrinos A. The Human Genome Project: lessons from large-scale biology. Science. 2003;300(5617):286–90.
17. Ashley EA. Towards precision medicine. Nat Rev Genet. 2016;17(9):507–22.
18. He M, Roussak K, Ma F, Borcherding N, Garin V, White M, Schutt C, Jensen TI, Zhao Y, Iberg CA, et al. CD5 expression by dendritic cells directs T cell immunity and sustains immunotherapy responses. Science. 2023;379(6633):eabg2752.
19. Garris CS, Arlauckas SP, Kohler RH, Trefny MP, Garren S, Piot C, Engblom C, Pfirschke C, Siwicki M, Gungabeesoon J, et al. Successful Anti-PD-1 Cancer Immunotherapy requires T cell-dendritic cell crosstalk involving the cytokines IFN-γ and IL-12. Immunity. 2018;49(6):1148–e11611147.
20. Chen Z, Huang Y, Hu Z, Zhao M, Li M, Bi G, Zheng Y, Liang J, Lu T, Jiang W, et al. Landscape and dynamics of single tumor and immune cells in early and advanced-stage lung adenocarcinoma. Clin Transl Med. 2021;11(3):e350.

21. Maynard A, McCoach CE, Rotow JK, Harris L, Haderk F, Kerr DL, Yu EA, Schenk EL, Tan W, Zee A, et al. Therapy-Induced Evolution of Human Lung Cancer revealed by single-cell RNA sequencing. Cell. 2020;182(5):1232–e12511222.
22. Kim N, Kim HK, Lee K, Hong Y, Cho JH, Choi JW, Lee JI, Suh YL, Ku BM, Eum HH, et al. Single-cell RNA sequencing demonstrates the molecular and cellular reprogramming of metastatic lung adenocarcinoma. Nat Commun. 2020;11(1):2285.
23. 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.
24. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. BMC Bioinformatics. 2008;9:559.
25. Liu Z, Liu L, Weng S, Guo C, Dang Q, Xu H, Wang L, Lu T, Zhang Y, Sun Z, et al. Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. Nat Commun. 2022;13(1):816.
26. Bastien P, Bertrand F, Meyer N, Maumy-Bertrand M. Deviance residuals-based sparse PLS and sparse kernel PLS regression for censored data. Bioinformatics. 2015;31(3):397–404.
27. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, Selves J, Laurent-Puig P, Sautès-Fridman C, Fridman WH, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. Genome Biol. 2016;17(1):218.
28. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. Genome Biol. 2017;18(1):220.
29. Yoshihara K, Shahmoradgoli M, Martínez E, Vegesna R, Kim H, Torres-Garcia W, Treviño V, Shen H, Laird PW, Levine DA, et al. Inferring tumour purity and stromal and immune cell admixture from expression data. Nat Commun. 2013;4:2612.
30. Song Y, Liu Y, Li ZM, Li L, Su H, Jin Z, Zuo X, Wu J, Zhou H, Li K, et al. SHR2554, an EZH2 inhibitor, in relapsed or refractory mature lymphoid neoplasms: a first-in-human, dose-escalation, dose-expansion, and clinical expansion phase 1 trial. Lancet Haematol. 2022;9(7):e493–503.
31. Li F, Ng WL, Luster TA, Hu H, Sviderskiy VO, Dowling CM, Hollinshead KER, Zouitine P, Zhang H, Huang Q, et al. Epigenetic CRISPR screens identify Npm1 as a therapeutic vulnerability in Non-small Cell Lung Cancer. Cancer Res. 2020;80(17):3556–67.
32. Li J, Yuan S, Norgard RJ, Yan F, Sun YH, Kim IK, Merrell AJ, Sela Y, Jiang Y, Bhanu NV, et al. Epigenetic and Transcriptional Control of the Epidermal Growth Factor Receptor regulates the Tumor Immune Microenvironment in Pancreatic Cancer. Cancer Discov. 2021;11(3):736–53.
33. Hinshaw DC, Shevde LA. The Tumor Microenvironment innately modulates Cancer Progression. Cancer Res. 2019;79(18):4557–66.
34. Ou Z, Lin S, Qiu J, Ding W, Ren P, Chen D, Wang J, Tong Y, Wu D, Chen A, et al. Single-nucleus RNA sequencing and spatial transcriptomics reveal the immunological microenvironment of cervical squamous cell carcinoma. Adv Sci (Weinh). 2022;9(29):e2203040.
35. Wang M, Cai Y, Peng Y, Xu B, Hui W, Jiang Y. Exosomal LGALS9 in the cerebrospinal fluid of glioblastoma patients suppressed dendritic cell antigen presentation and cytotoxic T-cell immunity. Cell Death Dis. 2020;11(10):896.
36. Abelin JG, Harjanto D, Malloy M, Suri P, Colson T, Goulding SP, Creech AL, Serrano LR, Nasir G, Nasrullah Y, et al. Defining HLA-II Ligand Processing and binding rules with Mass Spectrometry enhances Cancer Epitope Prediction. Immunity. 2019;51(4):766–e779717.
37. Grabowska J, Lopez-Venegas MA, Affandi AJ, den Haan JMM. CD169(+) macrophages capture and dendritic cells Instruct: the interplay of the gatekeeper and the General of the Immune System. Front Immunol. 2018;9:2472.
38. Heinz LX, Lee J, Kapoor U, Kartnig F, Sedlyarov V, Papakostas K, César-Razquin A, Essletzbichler P, Goldmann U, Stefanovic A, et al. TASL is the SLC15A4-associated adaptor for IRF5 activation by TLR7-9. Nature. 2020;581(7808):316–22.
39. Harris VM, Koelsch KA, Kurien BT, Harley ITW, Wren JD, Harley JB, Scofield RH. Characterization of cxorf21 provides Molecular Insight into Female-Bias Immune response in SLE pathogenesis. Front Immunol. 2019;10:2160.
40. Clark AM, Reynolds SH, Anderson M, Wiest JS. Mutational activation of the MAP3K8 protooncogene in lung cancer. Genes Chromosomes Cancer. 2004;41(2):99–108.
41. Shi X, Li R, Dong X, Chen AM, Liu X, Lu D, Feng S, Wang H, Cai K. IRGS: an immune-related gene classifier for lung adenocarcinoma prognosis. J Transl Med. 2020;18(1):55.
42. Xu YF, Yi Y, Qiu SJ, Gao Q, Li YW, Dai CX, Cai MY, Ju MJ, Zhou J, Zhang BH, et al. PEBP1 downregulation is associated to poor prognosis in HCC related to hepatitis B infection. J Hepatol. 2010;53(5):872–9.
43. Hu K, Qin X, Shao Y, Zhou Y, Ye G, Xu S. Circular RNA MTO1 suppresses tumorigenesis of gastric carcinoma by sponging mir-3200-5p and targeting PEBP1. Mol Cell Probes. 2020;52:101562.
44. Park S, Yeung ML, Beach S, Shields JM, Yeung KC. RKIP downregulates B-Raf kinase activity in melanoma cancer cells. Oncogene. 2005;24(21):3535–40.
45. Zou Q, Wu H, Fu F, Yi W, Pei L, Zhou M. RKIP suppresses the proliferation and metastasis of breast cancer cell lines through up-regulation of miR-185 targeting HMGA2. Arch Biochem Biophys. 2016;610:25–32.
46. Zhang L, Fu Z, Binkley C, Giordano T, Burant CF, Logsdon CD, Simeone DM. Raf kinase inhibitory protein inhibits beta-cell proliferation. Surgery. 2004;136(3):708–15.
47. Hellmann J, Rommelspacher H, Mühlbauer E, Wernicke C. Raf kinase inhibitor protein enhances neuronal differentiation in human SH-SY5Y cells. Dev Neurosci. 2010;32(1):33–46.
48. Baritaki S, Chapman A, Yeung K, Spandidos DA, Palladino M, Bonavida B. Inhibition of epithelial to mesenchymal transition in metastatic prostate cancer cells by the novel proteasome inhibitor, NPI-0052: pivotal roles of snail repression and RKIP induction. Oncogene. 2009;28(40):3573–85.
49. Khan I, Steeg PS. Metastasis suppressors: functional pathways. Lab Invest. 2018;98(2):198–210.
50. Frankenberger C, Rabe D, Bainer R, Sankarasharma D, Chada K, Krausz T, Gilad Y, Becker L, Rosner MR. Metastasis suppressors regulate the Tumor Microenvironment by blocking recruitment of Prometastatic Tumor-Associated macrophages. Cancer Res. 2015;75(19):4063–73.
51. Ma C, Li F, Luo H. Prognostic and immune implications of a novel ferroptosis-related ten-gene signature in lung adenocarcinoma. Ann Transl Med. 2021;9(13):1058.
52. Cheng T, Shan G, Yang H, Gu J, Lu C, Xu F, Ge D. Development of a ferroptosis-based model to predict prognosis, tumor microenvironment, and drug response for lung adenocarcinoma with weighted genes co-expression network analysis. Front Pharmacol. 2022;13:1072589.
53. Chatterjee D, Bai Y, Wang Z, Beach S, Mott S, Roy R, Braastad C, Sun Y, Mukhopadhyay A, Aggarwal BB, et al. RKIP sensitizes prostate and breast cancer cells to drug-induced apoptosis. J Biol Chem. 2004;279(17):17515–23.
54. Mondino A, Blasi F. uPA and uPAR in fibrinolysis, immunity and pathology. Trends Immunol. 2004;25(8):450–5.
55. Lund IK, Illemann M, Thurison T, Christensen IJ, Høyer-Hansen G. uPAR as anticancer target: evaluation of biomarker potential, histological localization, and antibody-based therapy. Curr Drug Targets. 2011;12(12):1744–60.
56. Ichiki K, Mitani N, Doki Y, Hara H, Misaki T, Saiki I. Regulation of activator protein-1 activity in the mediastinal lymph node metastasis of lung cancer. Clin Exp Metastasis. 2000;18(7):539–45.
57. Ye P, Zhao L, Gonda TJ. The MYB oncogene can suppress apoptosis in acute myeloid leukemia cells by transcriptional repression of DRAK2 expression. Leuk Res. 2013;37(5):595–601.
58. Szoltysek K, Ciardullo C, Zhou P, Walaszczyk A, Willmore E, Rand V, Marshall S, Hall A, C JH, Eswaran J et al. DAP kinase-related apoptosis-inducing protein kinase 2 (DRAK2) is a Key Regulator and molecular marker in chronic lymphocytic leukemia. Int J Mol Sci 2020, 21(20).
59. Doherty GA, Byrne SM, Austin SC, Scully GM, Sadlier DM, Neilan TG, Kay EW, Murray FE, Fitzgerald DJ. Regulation of the apoptosis-inducing kinase DRAK2 by cyclooxygenase-2 in colorectal cancer. Br J Cancer. 2009;101(3):483–91.
60. Shi X, Zhou Q, Huang B, Xia S, Jiang Y, Fang S, Lin J. Prognostic and immune-related value of STK17B in skin cutaneous melanoma. PLoS ONE. 2022;17(2):e0263311.
61. Lan Y, Han J, Wang Y, Wang J, Yang G, Li K, Song R, Zheng T, Liang Y, Pan S, et al. STK17B promotes carcinogenesis and metastasis via AKT/GSK-3β/Snail signaling in hepatocellular carcinoma. Cell Death Dis. 2018;9(2):236.
62. Yang KM, Kim W, Bae E, Gim J, Weist BM, Jung Y, Hyun JS, Hernandez JB, Leem SH, Park T, et al. DRAK2 participates in a negative feedback loop to control TGF-β/Smads signaling by binding to type I TGF-β receptor. Cell Rep. 2012;2(5):1286–99.
63. Dai J, Jiang M, He K, Wang H, Chen P, Guo H, Zhao W, Lu H, He Y, Zhou C. DNA damage response and repair gene alterations increase Tumor Mutational Burden and promote poor prognosis of Advanced Lung Cancer. Front Oncol. 2021;11:708294.
64. Han G, Yang G, Hao D, Lu Y, Thein K, Simpson BS, Chen J, Sun R, Alhalabi O, Wang R, et al. 9p21 loss confers a cold tumor immune microenvironment and primary resistance to immune checkpoint therapy. Nat Commun. 2021;12(1):5606.

65. Gerard CL, Delyon J, Wicky A, Homicsko K, Cuendet MA, Michielin O. Turning tumors from cold to inflamed to improve immunotherapy response. Cancer Treat Rev. 2021;101:102227.
66. Liu WJ, Du Y, Wen R, Yang M, Xu J. Drug resistance to targeted therapeutic strategies in non-small cell lung cancer. Pharmacol Ther. 2020;206:107438.
67. Zhang KR, Zhang YF, Lei HM, Tang YB, Ma CS, Lv QM, Wang SY, Lu LM, Shen Y, Chen HZ, et al. Targeting AKR1B1 inhibits glutathione de novo synthesis to overcome acquired resistance to EGFR-targeted therapy in lung cancer. Sci Transl Med. 2021;13(614):eabg6428.
68. Passaro A, Brahmer J, Antonia S, Mok T, Peters S. Managing resistance to Immune checkpoint inhibitors in Lung Cancer: treatment and novel strategies. J Clin Oncol. 2022;40(6):598–610.
69. Chen Q, Sun T, Jiang C. Recent advancements in Nanomedicine for 'Cold' Tumor Immunotherapy. Nanomicro Lett. 2021;13(1):92.
70. Li L, Sun R, Miao Y, Tran T, Adams L, Roscoe N, Xu B, Manyam GC, Tan X, Zhang H, et al. PD-1/PD-L1 expression and interaction by automated quantitative immunofluorescent analysis show adverse prognostic impact in patients with diffuse large B-cell lymphoma having T-cell infiltration: a study from the International DLBCL Consortium Program. Mod Pathol. 2019;32(6):741–54.
71. Scholl S, Beuzeboc P, Pouillart P. Targeting HER2 in other tumor types. Ann Oncol. 2001;12(Suppl 1):S81–87.
72. Yang K, Li C, Liu Y, Gu X, Jiang L, Shi L. Prognostic and immunotherapeutic roles of KRAS in Pan-cancer. Cells 2022, 11(9).

Publisher's note