

Original Article

Hypoxia-anoikis-related genes in LUAD: machine learning and RNA sequencing analysis of immune infiltration and therapy response

Yihao Liu^{1,2*}, Wenhao Zhao^{1,2*}, Zexia Zhao^{1,2*}, Zhixuan Duan^{1,2*}, Hua Huang^{1,2}, Chen Ding^{1,2}, Sensen Hou^{1,2}, Minghui Liu^{1,2}, Hongbing Zhang^{1,2}, Yongwen Li^{1,2}, Min Wang^{1,2}, Wenjun Meng³, Jun Chen^{1,2}, Haoling Zhang⁴, Honglin Zhao^{1,2}

¹Department of Lung Cancer Surgery, Tianjin Medical University General Hospital, Tianjin 300052, The People's Republic of China; ²Tianjin Key Laboratory of Lung Cancer Metastasis and Tumor Microenvironment, Tianjin Lung Cancer Institute, Tianjin Medical University General Hospital, Tianjin 300052, The People's Republic of China; ³Department of Pain Management, West China Hospital, Sichuan University, Chengdu 610041, Sichuan, The People's Republic of China; ⁴Department of Biomedical Sciences, Advanced Medical and Dental Institute, Universiti Sains Malaysia, Kepala Batas 13200, Penang, Malaysia. *Equal contributors.

Received July 1, 2025; Accepted August 25, 2025; Epub August 25, 2025; Published August 30, 2025

Abstract: Hypoxia plays a crucial role in the pathogenesis of various cancers, especially lung adenocarcinoma (LUAD), by altering cancer metabolism to promote escape mechanisms. Anoikis, a specialized form of programmed cell death, is evaded by LUAD cells during tumor progression and metastasis through upregulation of anti-apoptotic proteins. Investigating the impact of hypoxia-anoikis-related genes on prognosis and therapy prediction in LUAD is essential. Gene expression and clinical data from 489 LUAD patients and 49 normal tissues in The Cancer Genome Atlas (TCGA) dataset were used as the training set, while GSE72094, GSE31210, and GSE30219 datasets were used for validation. Weighted Gene Co-Expression Network Analysis (WGCNA) identified genes associated with hypoxia and anoikis. Machine learning models were evaluated using the C-index. Kaplan-Meier survival analysis, immune cell infiltration, tumor mutational burden (TMB), and sensitivity to therapy were assessed based on risk scores. A total of 21 hypoxia-anoikis-related prognostic genes were identified. The Random Survival Forest (RSF) model had the highest C-index. High-risk patients had significantly lower survival rates. Immune analysis showed higher immune infiltration in the low-risk group, with lower immune escape potential in these patients. Risk scores were correlated with sensitivity to targeted therapy and chemotherapy. MCF2 was identified as a key prognostic gene, and its knockdown inhibited LUAD cell proliferation and metastasis. These 21 genes offer insights into LUAD prognosis and therapy response, guiding personalized treatment strategies for LUAD patients.

Keywords: Hypoxia, anoikis, machine learning, biomarker detection, experimental validation of biomarker, targeted therapy, lung adenocarcinoma

Introduction

Lung adenocarcinoma (LUAD), recognized as both a highly prevalent and aggressive variant of non-small cell lung cancer (NSCLC), is distinguished by its significant metastatic capability and unfavorable prognosis [1]. Despite advancements in early detection and therapeutic interventions, the survival rate of LUAD patients continues to be poor, largely due to mechanisms such as immune system evasion by the tumor and resistance to programmed

cell death. Hypoxia, a common feature of solid malignancies, plays a critical role in tumor development, metabolic reprogramming, and resistance to therapy, particularly in LUAD. Under low-oxygen conditions, cancerous cells undergo both metabolic shifts and microenvironmental modifications, promoting immune escape while enhancing tumor survival and dissemination [2].

Anoikis, a specialized type of programmed cell death, takes place when cells lose appropriate

anchorage to the extracellular matrix (ECM) [3, 4]. This apoptotic process functions as a key defensive mechanism that prevents the spread of aberrant cells. However, malignant cells have developed multiple survival strategies to bypass anoikis, including: (1) increased activation of focal adhesion kinase (FAK) and Src kinase, which stimulate pro-survival pathways such as PI3K/Akt and MAPK/ERK1/2, alongside downregulation of pro-apoptotic Bcl-2 family proteins (e.g., Bax) and simultaneous upregulation of anti-apoptotic factors (e.g., Bcl-xL); (2) epithelial-mesenchymal transition (EMT)-driven changes in cell adhesion, marked by diminished E-cadherin expression and enhanced motility, allowing detached tumor cells to persist in circulation and facilitate metastasis; (3) modifications to the ECM that generate a tumor-supportive microenvironment, ultimately aiding cancer invasion and survival. These pathways are particularly relevant in LUAD progression.

Although the intricate interplay between hypoxia and anoikis resistance in LUAD remains only partially understood, current findings indicate a significant association. The hypoxic tumor microenvironment promotes survival by enabling mitochondrial translocation of phosphoglycerate kinase 1 (PGK1), initiating dual protective mechanisms against anoikis: (1) inhibition of pyruvate oxidation while fostering aerobic glycolysis (the Warburg effect) [5, 6], which decreases reactive oxygen species (ROS) levels and enhances lactate accumulation to sustain cellular energy homeostasis; (2) enhancement of the homing ability of disseminated malignant cells. This metabolic and transcriptional reprogramming, regulated by the hypoxia-PGK1 axis [7], not only grants survival advantages to LUAD cells that have detached but also heightens their likelihood of establishing distant metastases, such as peritoneal dissemination [8]. Previous studies have reported that hypoxia-inducible factor (HIF)-1 activation inhibits the expression of Bim and Bmf, reducing apoptosis and thereby preventing anoikis [9]. Additionally, ANGPTL4 enhances the cell's resistance to anoikis and promotes tumor growth and metastasis by activating the FAK/Src/PI3K-Akt/ERK signaling pathways [10]. Furthermore, HIF-1 α inhibits the expression of α 5 integrin, preventing anoikis in gastric cancer cells and promoting cell survival and metasta-

sis. HIF-1 α -deficient cells form fewer colonies in soft agar, indicating its importance in anchorage-independent proliferation [11].

Specifically, “hypoxia-anoikis-related genes” refer to genes that are activated in a hypoxic environment and inhibit or evade anoikis when tumor cells lose their attachment to the extracellular matrix (ECM). These genes help tumor cells escape anoikis by regulating mechanisms such as cell survival, metabolic reprogramming, and cell migration, thereby promoting tumor cell survival and metastasis. In this research, Weighted Gene Co-Expression Network Analysis (WGCNA) was utilized to pinpoint genes linked to both hypoxia and anoikis, followed by the application of 117 machine learning model combinations to select 21 prognostic genes [12-14]. A risk stratification model was developed to predict patient survival outcomes, response to targeted treatments, and chemotherapy sensitivity, thereby providing new insights into precision oncology and individualized treatment approaches.

Materials and methods

Data acquisition and processing

Gene expression and clinical data from 489 LUAD patients and 59 adjacent normal tissues were obtained from TCGA and used as the test cohort. External validation datasets (GSE72094, GSE31210, and GSE30219) were retrieved from the Gene Expression Omnibus (GEO). A total of 243 hypoxia-related genes were obtained from the Molecular Signatures Database (MSigDB), while 338 anoikis-related genes were curated from published literature ([Supplementary Table 1](#)) [15-17]. Single Sample Gene Set Enrichment Analysis (ssGSEA) was performed to calculate hypoxia and anoikis scores for each patient [18].

Weighted correlation network analysis (WGCNA) for hypoxia-anoikis gene identification

The WGCNA package in R was used to construct a gene co-expression network based on TCGA-LUAD transcriptomic data. WGCNA involves four core steps [19]: Calculation of gene-gene correlation coefficients. Identification of gene modules. Construction of a co-expression

network. Analysis of module-trait correlations. A soft threshold power (β) of 6 was selected to ensure network reliability, with a scale-free topology index exceeding 0.90. All mRNA genes were clustered into different modules based on their expression patterns, with different colors representing distinct modules. The correlation between these modules and hypoxia and anoikis scores was visualized using heatmaps and correlation coefficients, leading to the identification of core genes.

Differentially expressed genes (DEGs) and prognostic gene identification

Differential expression analysis was conducted using the “limma” R package with $|\log_2FC| > 2.0$ and $FDR < 0.01$ as cutoffs. Univariate Cox regression analysis was performed using the “survival” package to identify prognosis-related genes. DEGs and prognostic genes were visualized using EnhancedVolcano and forest plot visualization.

Machine learning-based hypoxia-anoikis risk score model

Ten machine learning algorithms (Lasso, RSF, StepCox, Elastic Net, Ridge, GBM, CoxBoost, plsRcox, SuperPC, survival-SVM) and 117 model combinations were used to construct the risk model [20]. The best model was selected based on the highest C-index. The model was validated in TCGA and external datasets (GSE72094, GSE31210, GSE30219). Risk scores were calculated using multivariate Cox regression, and patients were stratified into high- and low-risk groups.

Nomogram construction and validation

Univariate and multivariate Cox regression analyses identified independent prognostic factors (stage, T-stage, N-stage, risk score). A nomogram was developed using the “Regplot” R package, and its predictive accuracy was assessed using time-dependent ROC curves and calibration plots.

Tumor mutation burden (TMB) analysis

The “Maftools” R package was used to analyze gene mutation profiles in both high-risk and low-risk groups.

Immune infiltration analysis and prediction of immune therapy response

Immune-related pathway activity levels were analyzed using Single Sample Gene Set Enrichment Analysis (ssGSEA). Comparisons of immune expression levels between high- and low-risk groups were conducted to assess immune microenvironment differences. The Tumor Immune Dysfunction and Exclusion (TIDE) algorithm was applied to LUAD transcriptomic data to evaluate immune infiltration levels and immune dysfunction, predicting patients' responses to immunotherapy. A higher TIDE score indicates worse immunotherapy efficacy. The “pheatmap” R package was used to generate a correlation heatmap visualizing the relationships among 21 prognostic genes, 28 immune pathways, immune scores, and TIDE scores [21]. The Cibersort algorithm in combination with the “Xcell” package was applied to assess immune infiltration differences among patient groups and to profile the distribution of multiple immune cell subtypes.

Prediction of potentially sensitive drugs

The half-maximal inhibitory concentration (IC50) is a critical metric for assessing cellular drug sensitivity [22, 23]. IC50 values for each patient were calculated using the “oncoPredict” R package based on gene expression profiles. A total of 198 drugs were evaluated, and the top 30 experimental inhibitors and targeted small molecules with negative correlations were selected through Pearson correlation coefficient analysis. The top 9 drugs with the smallest correlation coefficients were chosen for further box plot analysis to compare sensitivity differences. Additionally, conventional targeted therapies and chemotherapy drugs were analyzed separately. The correlation between these drugs and risk scores was visualized using heatmaps.

Drug-gene network analysis

Cox regression analysis identified MCF2 as a high-risk prognostic gene, which was visualized using a forest plot. The DSigDB online database was used to retrieve gene-drug enrichment data. Further enrichment analysis was conducted to assess the association between MCF2 and potential therapeutic drugs. Sub-

sequently, a drug-MCF2 network was established to identify potential therapeutic agents for further investigation.

Single-cell RNA-seq data analysis

Previously published single-cell RNA sequencing dataset GSE127465 was downloaded from the GEO database [24], which includes sequencing data from seven patients with non-small cell lung cancer (NSCLC). The raw data were processed using the “Seurat” R package. Quality control steps included retaining cells with 500-3000 detected genes, excluding cells with > 10% mitochondrial gene content, and keeping genes expressed in at least three different cells. Cell type annotation was performed based on canonical biomarkers. The expression density of the MCF2 gene was examined across different cell clusters.

Cell culture and transfection

The A549 and H1299 human lung adenocarcinoma cell lines were obtained from the Cell Bank of the Chinese Academy of Sciences (Shanghai, China). Cells were maintained in RPMI-1640 medium (Gibco, USA) supplemented with 10% fetal bovine serum (FBS) at 37°C in a 5% CO₂ incubator. Small interfering RNA (siRNA) targeting MCF2 (si-MCF2) and a negative control siRNA (si-NC) were synthesized (Ribobio, China) for MCF2 knockdown. Transfection was performed using Lipofectamine 2000 (Invitrogen, Carlsbad, CA, USA) following the manufacturer's protocol. The siRNA primer sequences for MCF2 are as follows: 5'-GGAAUGUCAAGAAAGUAAA-3'; 5'-UAACUUU-CUUUGACAUUCCUG-3'.

Cell viability and proliferation assays

Cell viability was assessed using the Cell Counting Kit-8 (CCK-8) assay [23, 25]. After siRNA transfection (48-72 hours), 5000 cells per well were seeded into 96-well plates and incubated at 37°C with 5% CO₂ for 2, 24, 48, and 72 hours. Subsequently, 100 µL of a mixture containing 80 µL complete culture medium and 20 µL CCK-8 solution (K1018-5ml, APExBIO, Shanghai, China) was added to each well and incubated for 2 hours. Optical density was measured at 450 nm every 24 hours using a spectrophotometer. Each experiment was performed in triplicate.

Migration and invasion assays

The Transwell migration and invasion assays were conducted using H1299 and A549 cells transfected with si-MCF2 [26]. For the migration assay, 50,000 cells per well were suspended in serum-free medium and seeded in the upper chamber of Transwell inserts. The lower chamber was filled with 600 µL of 20% FBS-containing medium as a chemoattractant. For the invasion assay, a similar procedure was performed using Matrigel-coated inserts. Cells were incubated for 24 or 48 hours, followed by fixation with paraformaldehyde, staining with crystal violet, and washing with PBS three times [27, 28]. Images were captured, and cell counts were analyzed using GraphPad Prism software.

Statistical analysis

All experiments were performed at least three times. For comparisons between two groups, Wilcoxon rank-sum tests or Student's t-tests were used as appropriate. WGCNA was used to identify key gene modules. Univariate and multivariate Cox regression analyses were performed to identify prognostic genes. GraphPad Prism was used to generate PCR result visualizations. Statistical significance thresholds: *P < 0.05, **P < 0.01, ***P < 0.001, ****P < 0.0001. All analyses were performed using R software (version 4.3.5).

Databases and online tools used

(1) TCGA GDC: <http://cancergenome.nih.gov/>;
(2) GEO: <https://www.ncbi.nlm.nih.gov/geo/>; (3) TIDE: <http://tide.dfci.harvard.edu/>; (4) DSigDB: <http://dsigdb.tanlab.org/DSigDBv1.0/>.

Results

Identification of hypoxia-anoikis-related genes

In this study, we used WGCNA to identify hypoxia-anoikis-related gene modules. Analysis of TCGA-LUAD transcriptomic data led to the construction of a gene co-expression network and the identification of multiple gene modules correlated with clinical phenotypes. Network topology analysis (**Figure 1A**) demonstrated a strong correlation between different soft threshold power values (range: 1-20) and scale-free topology index. A soft threshold of $\beta = 6$

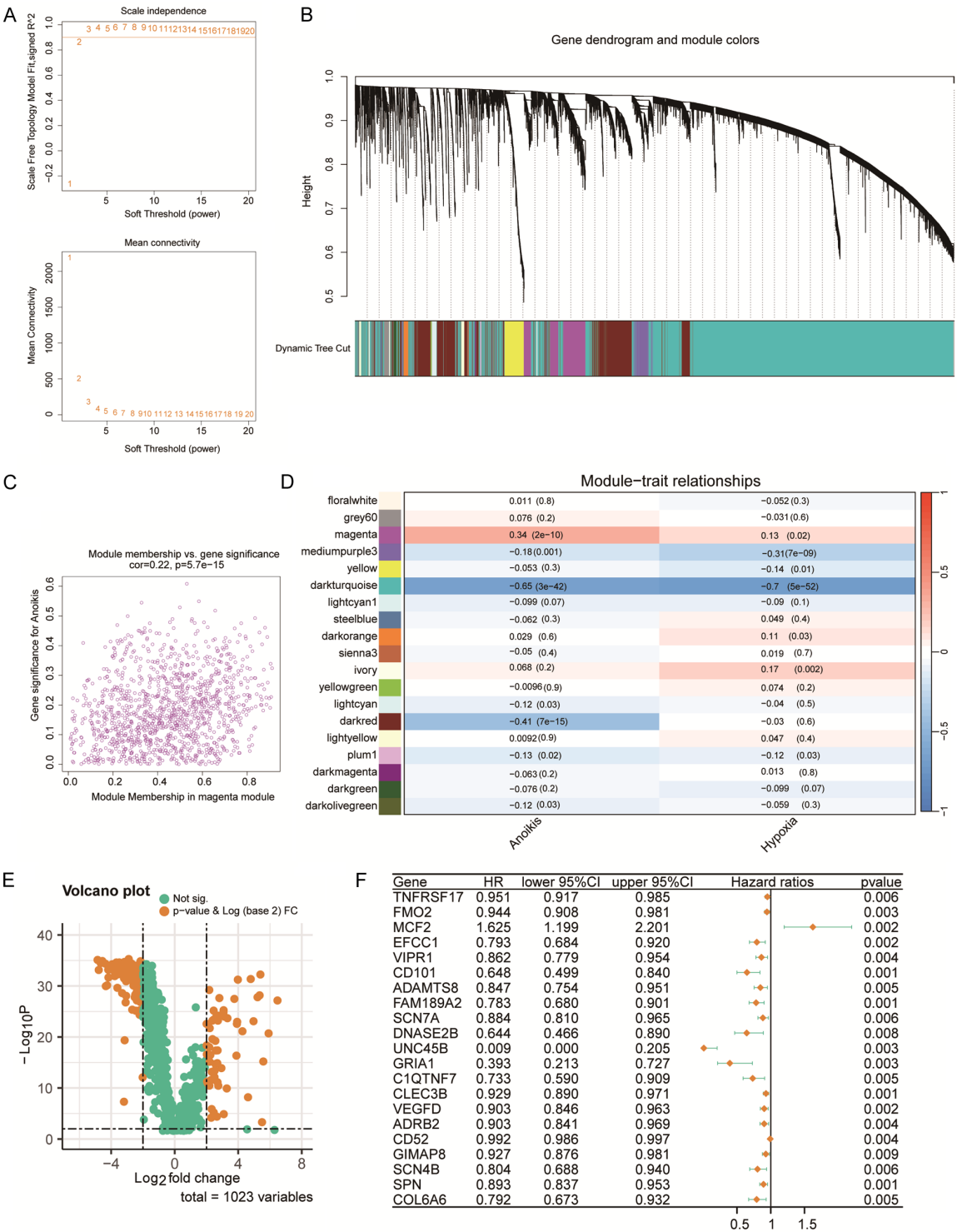


Figure 1. Construction of co-expression network and module-phenotype association analysis. A. Network topology analysis: Evaluating the scale-free topology of the network across different soft-thresholding powers (power range: 1-20). B. Gene module identification: A hierarchical clustering dendrogram based on the TOM illustrating gene co-expression relationships, with modules identified using the dynamic tree-cutting algorithm and color-coded. C. Scatter plot of genes in the magenta module. D. Module-phenotype association heatmap: Displaying correlation coefficients and p -values between gene modules and anoikis and hypoxia. E. Volcano plot showing differentially expressed genes in the module. F. Univariate Cox regression identifying 21 prognostic genes.

was selected to ensure network reliability. All mRNA genes were clustered into several modules based on hierarchical clustering of the Topological Overlap Matrix (TOM) (**Figure 1B**). Correlation heatmaps revealed significant associations between different gene modules and hypoxia-anoikis traits, with the magenta module showing the highest correlation with anoikis (correlation coefficient = 0.34) and a moderate correlation with hypoxia (correlation coefficient = 0.13). Overall, the magenta module genes exhibited the strongest correlation with both hypoxia and anoikis (**Figure 1C** and **1D**). The volcano plot (**Figure 1E**) identified significantly differentially expressed genes ($|\log_2\text{FC}| > 2.0$, $P < 0.01$). The magenta module genes and the list of differentially expressed genes are presented in [Supplementary Table 1](#). Univariate Cox analysis (**Figure 1F**) identified 21 genes significantly associated with LUAD prognosis, which may serve as potential prognostic biomarkers.

Construction and validation of the hypoxia-anoikis prognostic risk model

To construct a prognostic model based on hypoxia-anoikis-related genes, 489 LUAD patients from TCGA were used as the training cohort, while external validation was performed using GSE72094, GSE31210, and GSE30219 datasets. Ten machine learning algorithms and 117 model combinations were employed to determine key genes and compute risk scores using a 10-fold cross-validation framework. **Figure 2A** shows the performance of different machine learning models in the training dataset, with the random survival forest (RSF) algorithm achieving the highest C-index (0.726), making it the optimal model. RSF error plots indicated model stability (**Figure 2B**), with variable importance analysis highlighting key prognostic genes, particularly CLEC3B, FM02, ADRB2, CD101, and FAM189A2 (**Figure 2C**). Risk stratification using the established risk score model revealed significant differences in survival between high-risk and low-risk patients across TCGA, GSE72094, GSE31210, and GSE30219 datasets. Risk score distributions and survival outcomes were visualized (**Figure 2D-G**), demonstrating that low-risk patients consistently exhibited better OS.

Construction and validation of the nomogram

Univariate and multivariate Cox regression analyses identified clinical stage, T stage, N stage, and risk score as independent risk factors (**Figure 3A** and **3B**). A nomogram integrating these factors was developed to predict patient survival probabilities (**Figure 3C**). The ROC curves for 1-, 3-, 5-, and 10-year OS prediction showed AUC values of 0.7011, 0.7109, 0.7213, and 0.7207, respectively (**Figure 3D**). Calibration plots (**Figure 3E**) demonstrated good agreement between predicted and observed survival outcomes, supporting the reliability of the nomogram.

TMB and risk score association

TMB analysis revealed that mutation burden varied significantly between risk groups (**Figure 4**). The low-risk group (**Figure 4A**) exhibited a median mutation count of 111, while the high-risk group (**Figure 4B**) had a median of 233, indicating significantly higher TMB in high-risk patients. Mutation profiles of the top 20 most frequently mutated genes in each group were visualized (**Figure 4C** and **4D**). TP53, TTN, MUC16, RYR2, and LRP1B were the most frequently mutated genes in the low-risk group, while TP53, TTN, CSMD3, MUC16, and ZFH4 were dominant in the high-risk group. Notably, TP53 mutations were more prevalent in the high-risk group, suggesting a link between TP53 dysfunction and aggressive tumor behavior.

Immune infiltration and immune therapy response prediction

ssGSEA analysis revealed significant differences in 28 immune cell types between the high- and low-risk groups (**Figure 5A**). Low-risk patients exhibited higher levels of CD8⁺ T cells, CD4⁺ T cells, B cells, macrophages, and dendritic cells, suggesting a more active immune microenvironment.

Immune-related scores, including stromal score, immune score, and ESTIMATE score, were significantly higher in the low-risk group (**Figure 5B**). TIDE scores were lower in the low-risk group (**Figure 5C**), indicating reduced immune evasion potential. Heatmap analysis of immune-related gene expression (**Figure 5D**

Machine learning identifies hypoxia-anoikis genes for LUAD prognosis

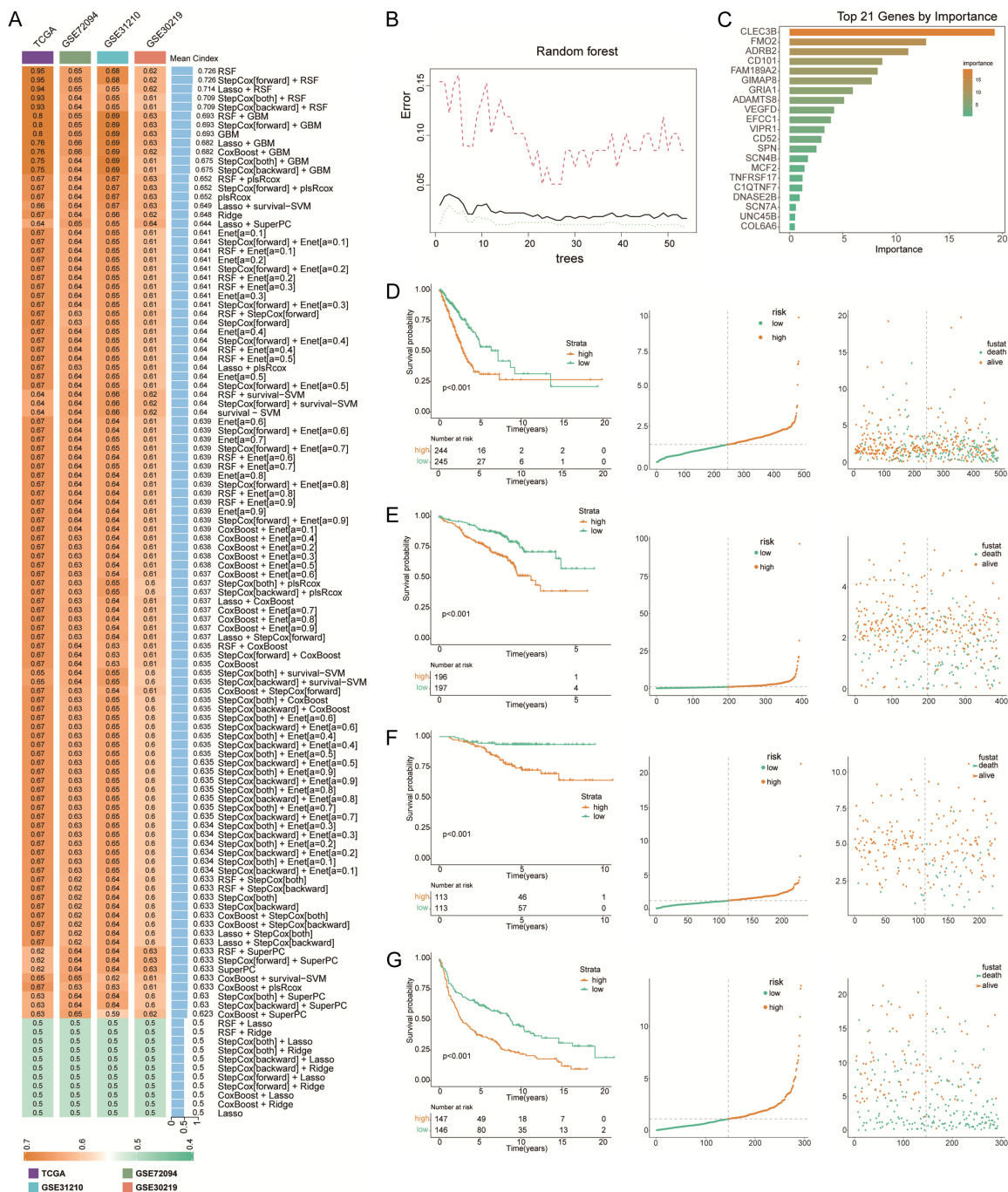


Figure 2. Prognostic genes identified by 117 machine learning models. **A.** Highest C-index across TCGA and three validation cohorts. **B.** Error plot of the RSF model. **C.** RSF feature importance ranking. **D.** Kaplan-Meier survival curve, survival status distribution, and risk score distribution for high- and low-risk groups in the TCGA test set. **E.** Kaplan-Meier survival curve, survival status distribution, and risk score distribution for high- and low-risk groups in the GSE72094 validation set. **F.** Kaplan-Meier survival curve, survival status distribution, and risk score distribution for high- and low-risk groups in the GSE31210 validation set. **G.** Kaplan-Meier survival curve, survival status distribution, and risk score distribution for high- and low-risk groups in the GSE30219 validation set.

and 5E) showed that genes associated with immune infiltration (e.g., TNFRSF17, CD52, SPN, and GIMAP8) were highly expressed in

low-risk patients. These findings suggest that low-risk LUAD patients may derive greater benefit from immune checkpoint inhibitor therapy,

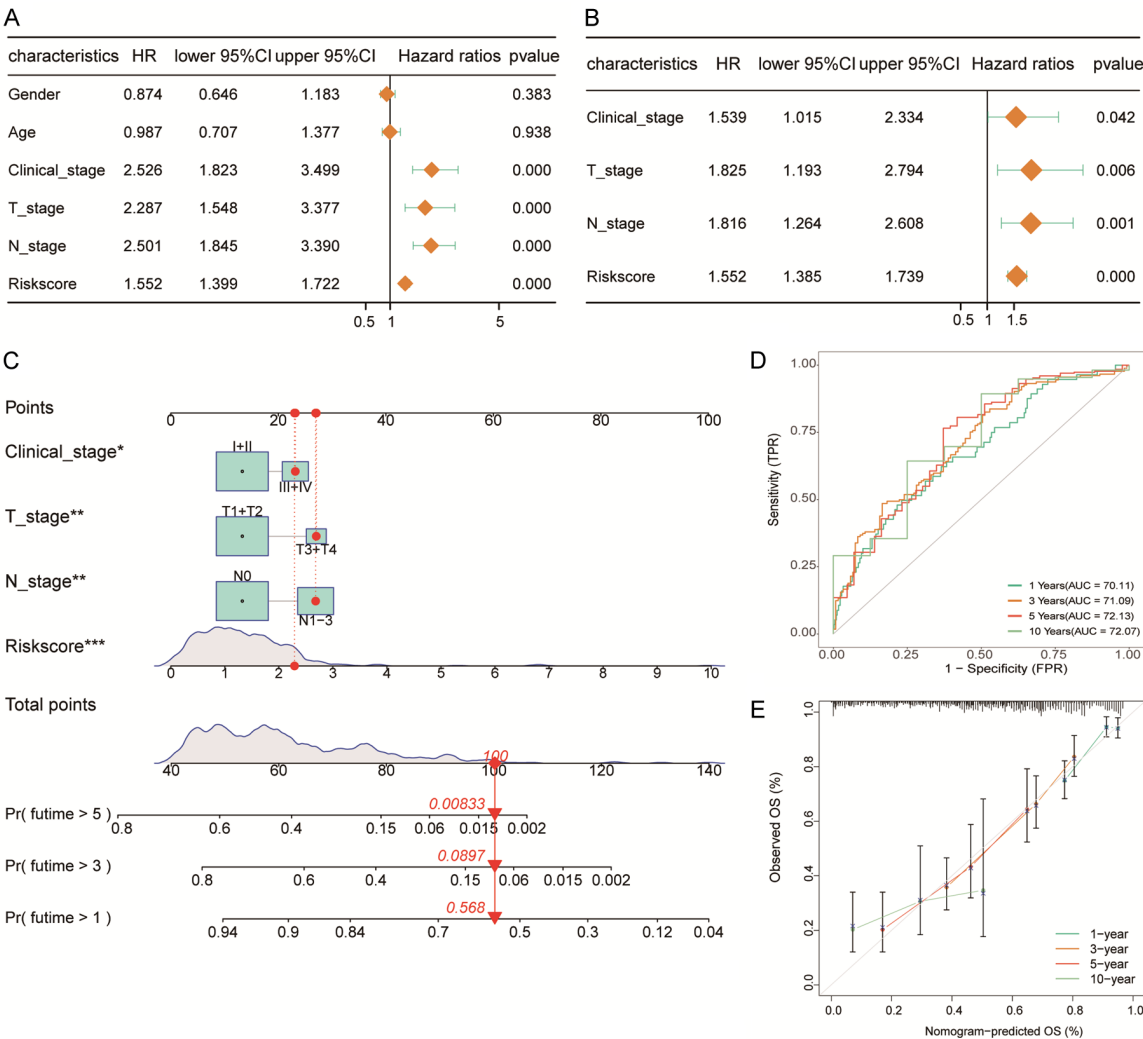


Figure 3. Development of a nomogram for predicting prognosis in LUAD patients. A. Univariate regression analysis. B. Multivariate Cox regression analysis identifying independent risk factors. C. Nomogram construction for predicting patient survival. D. ROC curves for TCGA patients at 1, 3, 5, and 10 years. E. Calibration curves comparing predicted and actual survival probabilities. *P < 0.05, **P < 0.01, ***P < 0.001.

whereas high-risk patients may exhibit immune resistance.

Model efficacy in predicting drug sensitivity

The IC50 values of the top 30 correlated therapy drugs were estimated using the oncoPredict R package (Figure 6A). High-risk patients exhibited significantly higher sensitivity to drugs including SCH772984, BI-2536, ERK_6604, AZD6738, AZD7762, MK-1775, Gallibiscoquinazole, LCL161, and AZD1332 (Figure 6B). Interestingly, in traditional chemotherapy and targeted therapy (Supplementary Figure 1A and 1B), the trend between risk scores and drug sensitivity differed. High-risk

patients may be insensitive to conventional chemotherapy drugs or show resistance, but exhibit certain sensitivity to some targeted drugs.

Analysis of MCF2 in drug response and immune microenvironment

MCF2 was identified as the only independent high-risk prognostic gene, with a hazard ratio (HR) of 1.886 (P < 0.001) (Figure 7A). Drug enrichment analysis revealed significant associations between MCF2 and Chromium(III) oxide, Chromium(II) chloride, and Cianidanol (p-values: 0.01, 0.02, and 0.04, respectively) (Figure 7B). The MCF2-drug interaction net-

Machine learning identifies hypoxia-anoikis genes for LUAD prognosis

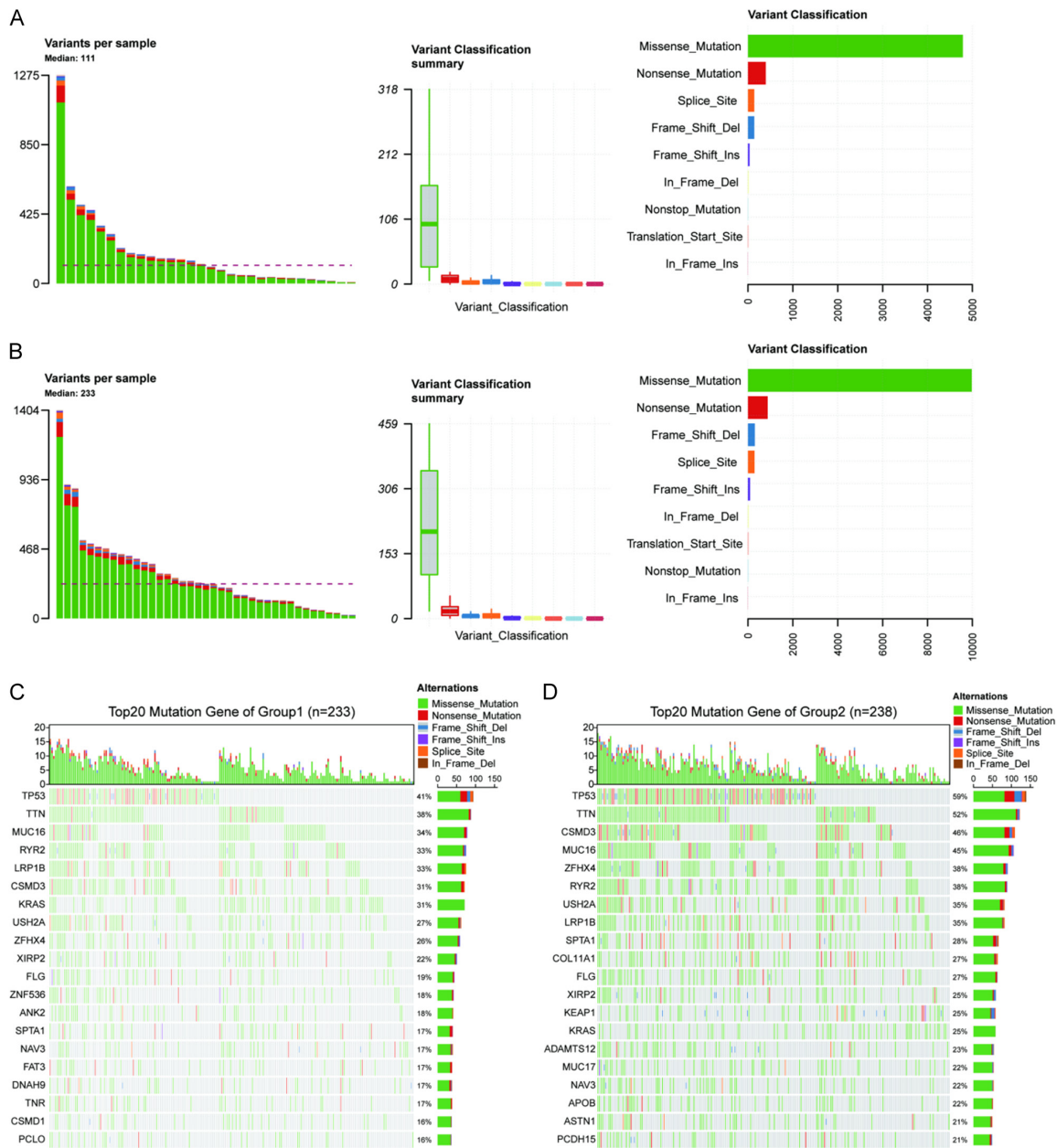


Figure 4. TMB in high- and low-risk groups. A. Number and types of gene mutations in the low-risk group. B. Number and types of gene mutations in the high-risk group. C. Top 20 mutated genes in the low-risk group. D. Top 20 mutated genes in the high-risk group.

work is shown in **Figure 7C**, suggesting that MCF2 may be a novel therapeutic target. Patients were divided into high and low MCF2 expression groups. Immune scores were calculated, revealing that both stromal score and immune score were higher in patients with elevated MCF2 expression (**Figure 7D**). CIBERSORT analysis indicated differences in immune cell infiltration proportions between groups

(**Figure 7E**). Using the Xcell algorithm, patients with high MCF2 expression exhibited increased abundance of immune cell subsets such as aDCs, B cells, and cDCs, while subsets such as Th2 cells were decreased (**Figure 7F**). At the single-cell level, MCF2 expression was predominantly enriched in NK cells and CD4⁺ Tn cells (**Figure 7G and 7H**). Taken together, these findings suggest that MCF2 influences both

Machine learning identifies hypoxia-anoikis genes for LUAD prognosis

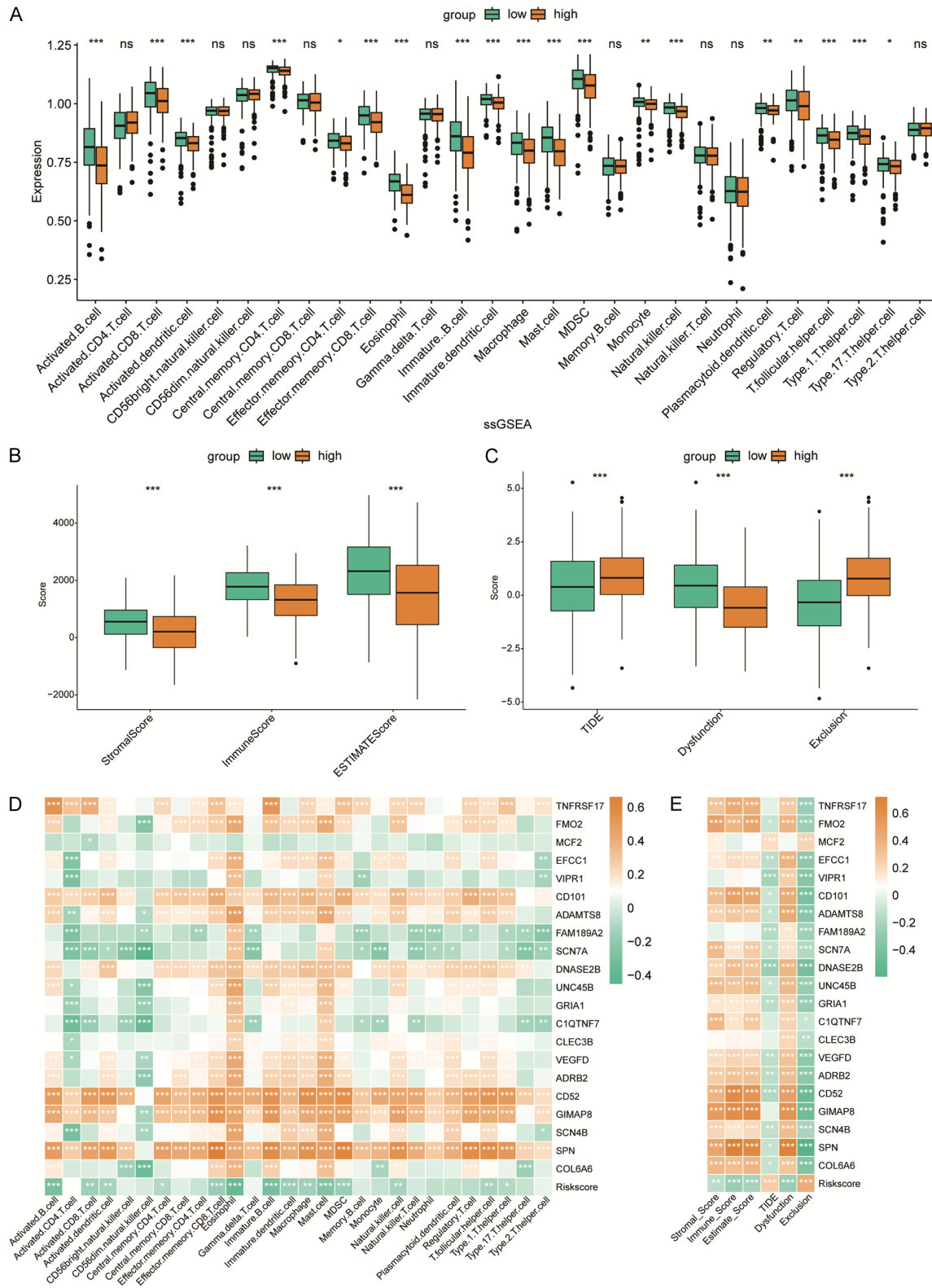
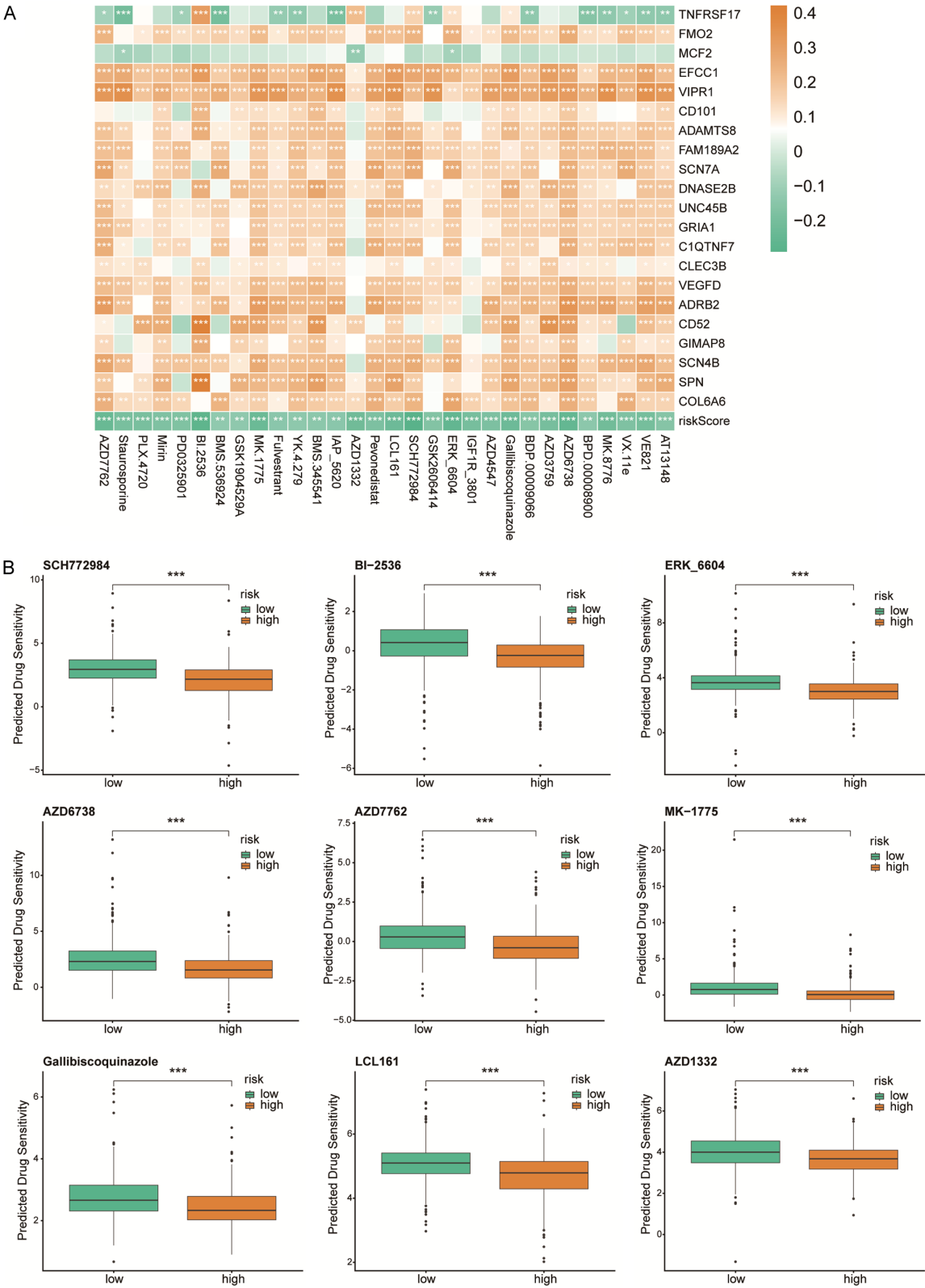


Figure 5. TME and response to immunotherapy. A. Expression levels of 28 immune cell types in high- and low-risk groups based on the ssGSEA algorithm. B. Immune score. C. TIDE (Tumor Immune Dysfunction and Exclusion) score. D. Heatmap of correlations between 21 prognostic genes and 28 immune cell types. E. Heatmap of correlations between 21 prognostic genes, risk score, immune score, and TIDE score. *P < 0.05, **P < 0.01, ***P < 0.001.



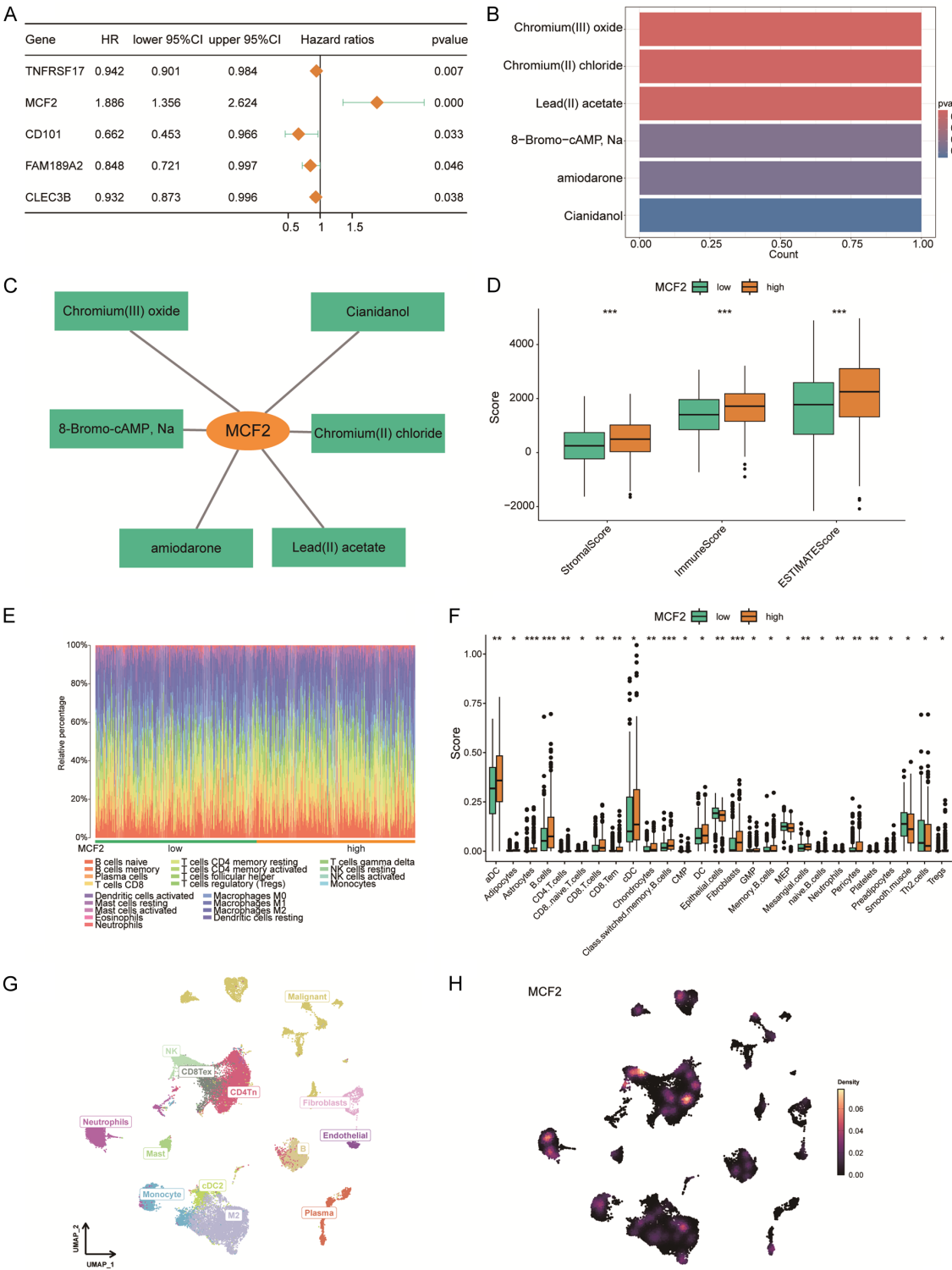


Figure 7. MCF2 of drug network analysis and immune microenvironment. A. Prognostic genes identified through multivariate Cox regression analysis. B. Analysis of associations between different drugs and the MCF2 gene. C. Molecular drug network illustrating potential pathways of multiple drugs targeting the MCF2 gene. D. MCF2 of ESTIMATE scores. E. Cibersort proportion plot showing the distribution of 21 immune cell types in MCF2 expression subgroups. F. Xcell boxplot showing the expression of immune cell types in MCF2 expression subgroups. G. Umap visualization of different immune cell types. H. Expression of MCF2 in immune cells at the single-cell level.

immune pathways and immune cell composition, and that changes in MCF2 expression may affect the regulation of tumor progression or invasion by NK cells and CD4⁺ T cells.

Survival analysis of LUAD patients with different clinical characteristics based on risk stratification

Kaplan-Meier survival analysis showed that, across different clinical characteristics, the survival probability of LUAD patients was significantly affected by risk scores. For patients aged ≥ 65 years, the high-risk group had significantly worse survival (**Figure 8A and 8B**). Male patients in the high-risk group showed poorer prognosis ($P = 0.002$), while female patients in the high-risk group also exhibited lower survival probability ($P = 0.001$) (**Figure 8C and 8D**). Regarding tumor staging, T1+T2 stage patients in the low-risk group had significantly better survival ($P < 0.001$), whereas T3+T4 stage patients in the high-risk group had poorer prognosis ($P = 0.013$) (**Figure 8E and 8F**). In terms of lymph node status, N0 stage patients in the low-risk group had significantly better survival ($P = 0.002$), while N1-3 stage patients did not show a significant difference in survival between the high- and low-risk groups ($P = 0.060$) (**Figure 8G and 8H**). For distant metastasis, M0 stage patients in the low-risk group had significantly better survival ($P < 0.001$), whereas M1 stage patients did not show a significant survival difference ($P = 0.484$) (**Figure 8I and 8J**). In clinical staging, I-II stage patients in the low-risk group exhibited significantly better survival ($P < 0.001$), while III-IV stage patients did not show a significant difference ($P = 0.106$) (**Figure 8K and 8L**). Overall, patients in the low-risk group demonstrated better survival prognosis across most clinical subgroups, but some groups (such as M1 and N1-3 stages) did not show a significant difference in risk score stratification.

MCF2 knockdown suppresses LUAD cell proliferation and invasion

MCF2 knockdown experiments were conducted in H1299 and A549 LUAD cell lines. CCK-8 assays demonstrated a significant reduction in cell proliferation following MCF2 silencing (**Figure 9A and 9B**). Transwell assays revealed that MCF2 knockdown significantly inhibited

cell migration and invasion (**Figure 9C and 9D**). These findings suggest that MCF2 plays a key role in promoting LUAD progression.

Discussion

This research primarily focuses on examining the combined value of genes related to hypoxia and anoikis in predicting the prognosis and guiding targeted chemotherapy for LUAD. Using WGCNA, we identified a range of gene modules linked to both hypoxia and anoikis, selecting the most strongly correlated group for further study. Subsequently, we developed the best risk score model by evaluating 117 different machine learning techniques, which categorized patients into high- and low-risk groups. The ssGSEA method was then used to assess and compare immune cell infiltration, survival predictions, tumor genetic mutations, and responses to targeted and chemotherapy drugs between these two groups. Among the genes with independent prognostic significance, MCF2 emerged as a key risk factor, and a drug-related molecular network for this gene was mapped. Our findings revealed that patients in the high-risk group had generally shorter survival rates compared to those in the low-risk group, though they exhibited greater sensitivity to most drugs. Thus, this risk model may play a pivotal role in clinical decision-making regarding prognosis and therapeutic strategies for LUAD.

HIF, a crucial transcription factor in the response to hypoxia, comprises several subunits that are essential for processes such as energy intake, tumor metabolism, cell proliferation and apoptosis, and adaptation to low-oxygen conditions. Under hypoxic conditions, LUAD cells activate genes associated with angiogenesis, like VEGF-A, while simultaneously inhibiting ECM components and CX26 connexins, facilitating metastasis and invasion [29]. Additionally, hypoxic tumor cells release more exosomes, which enhance intercellular communication and modify tumor cell metabolism and growth. The M2 macrophage phenotype, which suppresses immune responses and aids tumor growth, is promoted under hypoxic conditions, further enhancing macrophage M2 expression [30]. Hypoxia also drives tumor cells to utilize the Warburg effect, converting glucose into lactate, which not only creates an acidic environment favorable for tumor growth but also optimizes energy production.

Machine learning identifies hypoxia-anoikis genes for LUAD prognosis

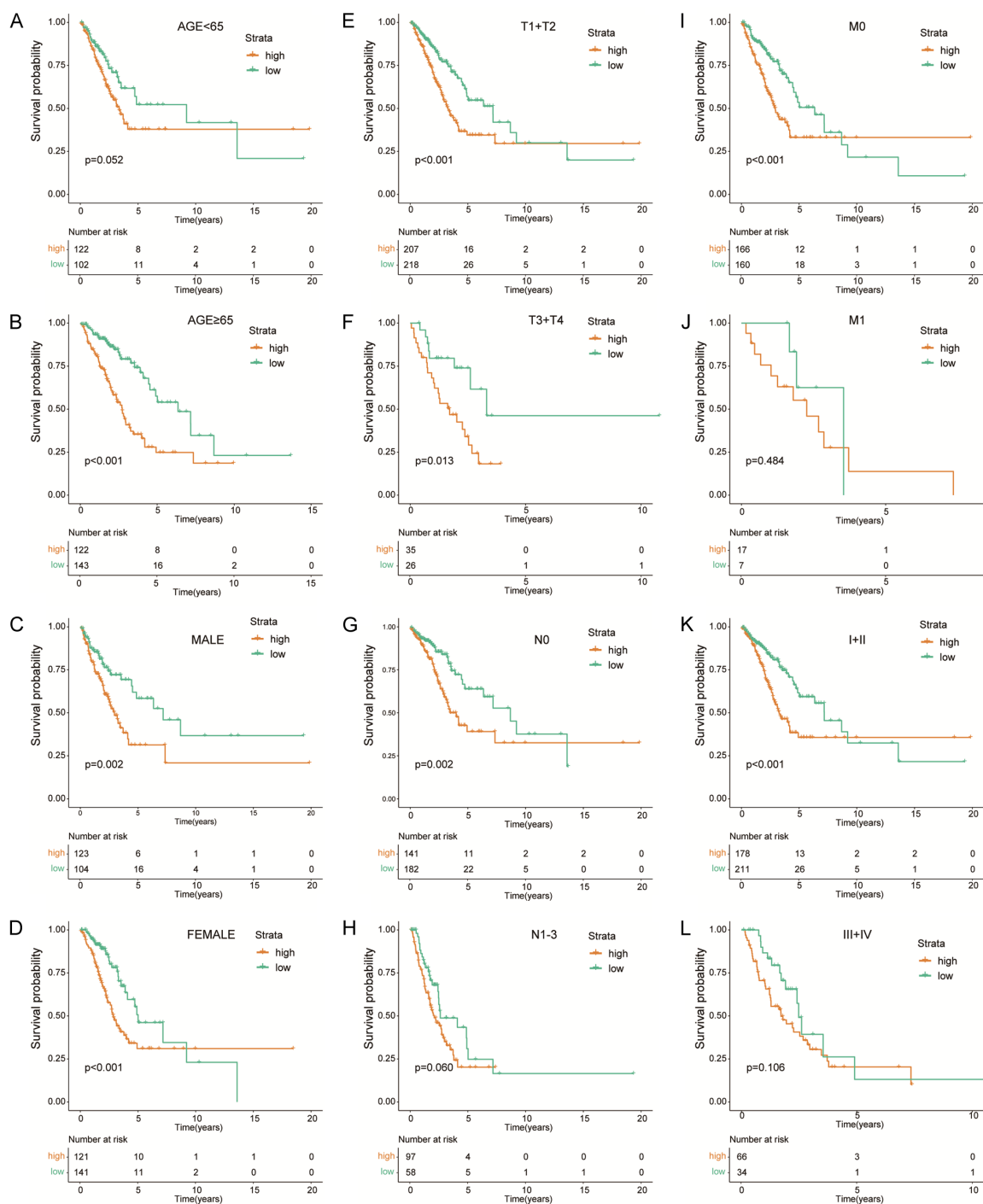


Figure 8. Kaplan-Meier curves for high- and low-risk groups. A. Kaplan-Meier curves for patients aged < 65. B. Kaplan-Meier curves for patients aged ≥ 65. C. Kaplan-Meier curves for male patients. D. Kaplan-Meier curves for female patients. E. Kaplan-Meier curves for patients in T1+T2 stages. F. Kaplan-Meier curves for patients in T3+T4 stages. G. Kaplan-Meier curves for patients in N0 stage. H. Kaplan-Meier curves for patients in N1-3 stages. I. Kaplan-Meier curves for patients in M0 stage. J. Kaplan-Meier curves for patients in M1 stage. K. Kaplan-Meier curves for patients in clinical stage I+II. L. Kaplan-Meier curves for patients in clinical stage III+IV.

Furthermore, hypoxia promotes the accumulation of LC stem cells, triggering EMT and

increasing resistance to drugs like gefitinib and ALK inhibitors [31].

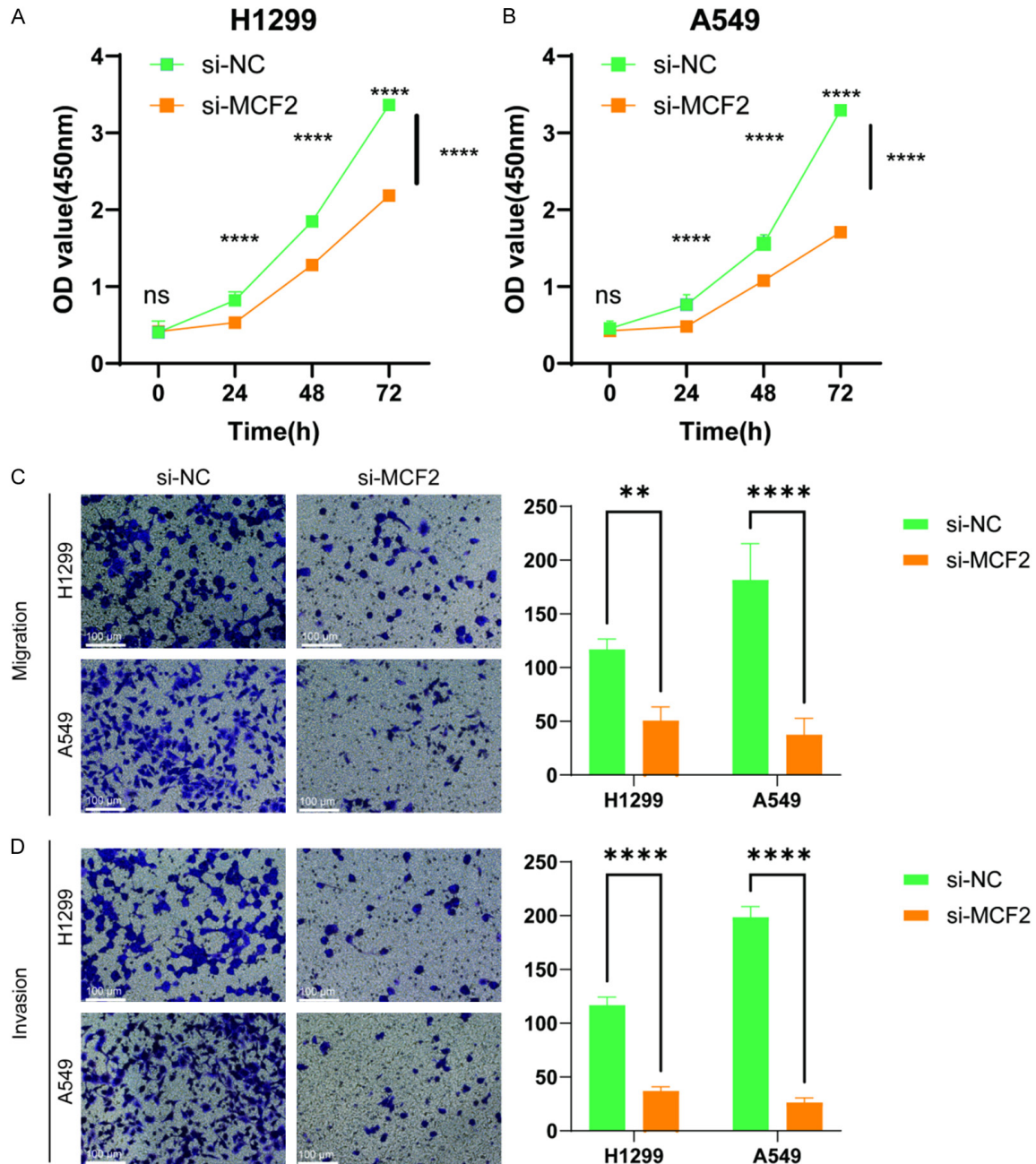


Figure 9. Biological function of MCF2 in LUAD cell lines. A. CCK-8 assay showing the effect of MCF2 knockdown on H1299 cell proliferation. B. CCK-8 assay showing the effect of MCF2 knockdown on A549 cell proliferation. C. Comparison of migration abilities between the NC group and the MCF2 knockdown group in H1299 and A549 cells. D. Comparison of invasion abilities between the NC group and the MCF2 knockdown group in H1299 and A549 cells. ** $P < 0.01$, **** $P < 0.0001$.

A critical factor in the distant metastasis of tumor cells is the disruption of apoptotic or regulated cell death pathways. Anoikis, a specialized form of programmed cell death, plays a vital role in preventing distant metastasis in tumors. Resistance to anoikis is a key mecha-

nism behind tumor progression [32, 33]. This resistance involves the activation of caspases, nucleases, and DNA fragmentation, which typically result in cell death. However, anoikis resistance can arise from multiple mechanisms, including the overexpression of growth factor

receptors in tumor cells. In NSCLC, anoikis resistance and metastasis are often driven by the upregulation of EMT, which also enhances the production of pleural effusion and promotes cancer cell invasion and metastasis [34]. Several intracellular metabolites, such as GDH1 and P-cadherin, are known to influence anoikis resistance, further facilitating tumor spread [35]. While no studies have conclusively established the connection between hypoxia and anoikis, identifying genes that link these two processes is crucial for treatment strategies.

The optimal model for predicting prognosis was developed using RSF machine learning, with the highest C-index value of 0.726. Through this model, 21 prognostic genes were identified, which are critical for patient survival and treatment decisions. ROC curve analysis indicated that the AUC values for 1-, 3-, 5-, and 10-year survival predictions ranged between 0.70 and 0.72, and Kaplan-Meier survival analysis showed that patients in the high-risk group had significantly shorter survival compared to those in the low-risk group. These 21 genes have demonstrated substantial potential in previous studies to affect tumor cell behavior.

In research by Dr. Zhao, VIPR1, the vasoactive intestinal peptide receptor-1, was shown to inhibit the progression of human lung adenocarcinoma cells [36]. Overexpressing VIPR1 suppressed cell growth, migration, and invasion in H1299 cells. Dr. Fu's findings further suggested that VIPR1 inhibits pyrimidine synthesis by regulating arginine, thereby preventing liver cancer development [37]. Research by Dr. Zhang revealed that ADAMTS8, an enzyme with thrombospondin motifs, could suppress breast cancer cell proliferation and invasion while inducing apoptosis [38]. VEGFD, when inhibited by S-nitrosylation, was found to promote LUAD progression [39]. Dr. Cheon's work highlighted the role of NPR2 as a co-receptor for VEGFR, where its inhibition disrupts signaling pathways, leading to fibroblast deactivation and preventing LUAD metastasis [40]. Additionally, COL6A6, essential for maintaining cellular structure, was shown to inhibit NSCLC proliferation via the JAK signaling pathway and could serve as a potential biomarker for prognosis and immune therapy [41]. Dr. Lu's research indicated that CLEC3B, involved

in the P53 signaling pathway, when expressed at low levels, can enhance LUAD metastasis, migration, and EMT [42].

Our findings revealed significant differences in immune and stromal scores between high- and low-risk groups. Furthermore, variations in immune pathways and cell expression were observed. The interaction between immune responses, hypoxia, and anoikis plays a key role in tumor progression, as hypoxia induces the Warburg effect, suppresses immune responses, and thereby promotes tumor metastasis and invasion. This underscores the importance of our prognostic model, which analyzes these relationships and aids in predicting responses to immune and targeted therapies.

Abemaciclib, a targeted therapy, has been shown to inhibit thyroid cancer's resistance to anoikis [43]. Similarly, AKT and MEK inhibitors, along with multi-tyrosine kinase inhibitors, can counteract anoikis resistance. Nanomedicine also holds promise, with human serum albumin nanoparticles (tHSA-NPs) exhibiting the ability to target and inhibit metastatic anoikis-resistant tumor cells. In our predictions, several chemotherapy and targeted drugs demonstrated sensitivity, particularly among high-risk patients [44, 45]. These drugs could serve as important indicators for evaluating patient prognosis.

While our model presents a promising tool, it does have limitations, including the lack of validation with a larger clinical sample. Moreover, as this study is retrospective, further prospective studies are necessary. The nomogram could also incorporate additional clinical factors, such as choices regarding adjuvant therapy and surgical intervention. Therefore, this model requires further investigation, and we hope that future research will validate and expand the hypoxia-anoikis-related prognostic model using diverse methodologies.

Conclusion

This study provides a broad and multi-dimensional perspective on the involvement of hypoxia-anoikis-related genes in LUAD prognosis and treatment decision-making. The risk score model serves a dual function as both a personalized prognostic framework and a molecular tool for precision oncology, linking gene expres-

sion data with drug response predictions. Furthermore, MCF2 has emerged as a critical oncogenic factor and a promising therapeutic target, offering a new avenue for future research. By enhancing our understanding of LUAD at a molecular level, these findings may facilitate the development of novel treatment strategies, ultimately contributing to the advancement of personalized medicine for lung cancer patients.

Acknowledgements

This research was funded by the Tianjin Municipal Health Commission, the Tianjin Key Medical Discipline Sub-project (TJLCMS2021-06), the Tianjin Municipal Education Commission through the General Project of the Natural Science Foundation (2020KJ-162), the Wu Jieping Medical Foundation (320.6750.2022-11-43), the National Natural Science Foundation of China (82172569), the Natural Science Foundation of Tianjin (23JCYBJC01010), and the Tianjin Key Medical Discipline (Specialty) Construction Project (TJYXZDXK-061B). Additional support was provided by the Key R&D Program for International Science and Technology Cooperation of the Gansu Provincial Science and Technology Program (23YFWA-0005), the Innovation Fund for Colleges and Universities of Gansu Province (2025A-110), the Collaborative Innovation Center for Traditional Chinese Medicine Prevention and Control of Nutrition- and Environment-Related Diseases in Northwest China (ZYXT-24-02), the Natural Science Project of the Gansu Provincial Department of Science and Technology (22JRSRA582), and the Natural Science Project of the Gansu Provincial Department of Science and Technology (2023A-088).

Disclosure of conflict of interest

None.

Address correspondence to: Haoling Zhang, Department of Biomedical Sciences, Advanced Medical and Dental Institute, Universiti Sains Malaysia, Kepala Batas 13200, Penang, Malaysia. E-mail: zhanghaolingedu@163.com; Honglin Zhao and Jun Chen, Department of Lung Cancer Surgery, Tianjin Medical University General Hospital, Tianjin 300052, The People's Republic of China. E-mail: navyzhao@aliyun.com (HLZ); huntercj2004@qq.com (JC)

References

- [1] Cords L, Engler S, Haberecker M, Ruschoff JH, Moch H, de Souza N and Bodenmiller B. Cancer-associated fibroblast phenotypes are associated with patient outcome in non-small cell lung cancer. *Cancer Cell* 2024; 42: 396-412, e395.
- [2] Yan F, Teng Y, Li X, Zhong Y, Li C, Yan F and He X. Hypoxia promotes non-small cell lung cancer cell stemness, migration, and invasion via promoting glycolysis by lactylation of SOX9. *Cancer Biol Ther* 2024; 25: 2304161.
- [3] Wang J, Luo Z, Lin L, Sui X, Yu L, Xu C, Zhang R, Zhao Z, Zhu Q, An B, Wang Q, Chen B, Leung EL and Wu Q. Anoikis-associated lung cancer metastasis: mechanisms and therapies. *Cancers (Basel)* 2022; 14: 4791.
- [4] Dai Y, Zhang X, Ou Y, Zou L, Zhang D, Yang Q, Qin Y, Du X, Li W, Yuan Z, Xiao Z and Wen Q. Anoikis resistance-protagonists of breast cancer cells survive and metastasize after ECM detachment. *Cell Commun Signal* 2023; 21: 190.
- [5] Duan W, Liu W, Xia S, Zhou Y, Tang M, Xu M, Lin M, Li X and Wang Q. Warburg effect enhanced by AKR1B10 promotes acquired resistance to pemetrexed in lung cancer-derived brain metastasis. *J Transl Med* 2023; 21: 547.
- [6] Chen X, Hao B, Li D, Reiter RJ, Bai Y, Abay B, Chen G, Lin S, Zheng T, Ren Y, Xu X, Li M and Fan L. Melatonin inhibits lung cancer development by reversing the Warburg effect via stimulating the SIRT3/PDH axis. *J Pineal Res* 2021; 71: e12755.
- [7] Guo Z, Zhang Y, Wang H, Liao L, Ma L, Zhao Y, Yang R, Li X, Niu J, Chu Q, Fu Y, Li B and Yang C. Hypoxia-induced downregulation of PGK1 crotonylation promotes tumorigenesis by coordinating glycolysis and the TCA cycle. *Nat Commun* 2024; 15: 6915.
- [8] Wilson RB, Solass W, Archid R, Weinreich FJ, Konigsrainer A and Reymond MA. Resistance to anoikis in transcoelomic shedding: the role of glycolytic enzymes. *Pleura Peritoneum* 2019; 4: 20190003.
- [9] Whelan KA, Caldwell SA, Shahriari KS, Jackson SR, Franchetti LD, Johannes GJ and Reginato MJ. Hypoxia suppression of Bim and Bmf blocks anoikis and luminal clearing during mammary morphogenesis. *Mol Biol Cell* 2010; 21: 3829-3837.
- [10] Baba K, Kitajima Y, Miyake S, Nakamura J, Wakiyama K, Sato H, Okuyama K, Kitagawa H, Tanaka T, Hiraki M, Yanagihara K and Noshiro H. Hypoxia-induced ANGPTL4 sustains tumour growth and anoikis resistance through different mechanisms in scirrhous gastric cancer cell lines. *Sci Rep* 2017; 7: 11127.

- [11] Rohwer N, Welzel M, Daskalow K, Pfander D, Wiedenmann B, Detjen K and Cramer T. Hypoxia-inducible factor 1 α mediates anoikis resistance via suppression of α 5 integrin. *Cancer Res* 2008; 68: 10113-10120.
- [12] Chen Y, Huang W, Ouyang J, Wang J and Xie Z. Identification of anoikis-related subgroups and prognosis model in liver hepatocellular carcinoma. *Int J Mol Sci* 2023; 24: 2862.
- [13] Qin Y, Pu X, Hu D and Yang M. Machine learning-based biomarker screening for acute myeloid leukemia prognosis and therapy from diverse cell-death patterns. *Sci Rep* 2024; 14: 17874.
- [14] Zhu J, Zhang J, Lou Y, Zheng Y, Zheng X, Cen W, Ye L and Zhang Q. Developing a machine learning-based prognosis and immunotherapeutic response signature in colorectal cancer: insights from ferroptosis, fatty acid dynamics, and the tumor microenvironment. *Front Immunol* 2024; 15: 1416443.
- [15] Mo G, Long X, Hu Z, Tang Y and Zhou Z. Anoikis-related gene signatures can aid prognosis of lung adenocarcinoma. *Adv Clin Exp Med* 2024; 33: 751-761.
- [16] Hu G, Li J, Zeng Y, Liu L, Yu Z, Qi X, Liu K and Yao H. The anoikis-related gene signature predicts survival accurately in colon adenocarcinoma. *Sci Rep* 2023; 13: 13919.
- [17] Jiang G, Song C, Wang X, Xu Y, Li H, He Z, Cai Y, Zheng M and Mao W. The multi-omics analysis identifies a novel cuproptosis-anoikis-related gene signature in prognosis and immune infiltration characterization of lung adenocarcinoma. *Heliyon* 2023; 9: e14091.
- [18] Chen Y, Feng Y, Yan F, Zhao Y, Zhao H and Guo Y. A novel immune-related gene signature to identify the tumor microenvironment and prognose disease among patients with oral squamous cell carcinoma patients using ssGSEA: a bioinformatics and biological validation study. *Front Immunol* 2022; 13: 922195.
- [19] Xu M, Zhou H, Hu P, Pan Y, Wang S, Liu L and Liu X. Identification and validation of immune and oxidative stress-related diagnostic markers for diabetic nephropathy by WGCNA and machine learning. *Front Immunol* 2023; 14: 1084531.
- [20] Cui Y, Zhang W, Lu W, Feng Y, Wu X, Zhuo Z, Zhang D and Zhang Y. An exosome-derived lncRNA signature identified by machine learning associated with prognosis and biomarkers for immunotherapy in ovarian cancer. *Front Immunol* 2024; 15: 1228235.
- [21] Huang X, Liu Y, Qian C, Shen Q, Wu M, Zhu B and Feng Y. CHSY3 promotes proliferation and migration in gastric cancer and is associated with immune infiltration. *J Transl Med* 2023; 21: 474.
- [22] Wang Y, Bin T, Tang J, Xu XJ, Lin C, Lu B and Sun TT. Construction of an acute myeloid leukemia prognostic model based on m6A-related efferocytosis-related genes. *Front Immunol* 2023; 14: 1268090.
- [23] Wu Z, Lu Z, Li L, Ma M, Long F, Wu R, Huang L, Chou J, Yang K, Zhang Y, Li X, Hu G, Zhang Y and Lin C. Identification and validation of ferroptosis-related lncRNA signatures as a novel prognostic model for colon cancer. *Front Immunol* 2022; 12: 783362.
- [24] Zilionis R, Engblom C, Pfirschke C, Savova V, Zemmour D, Saatcioglu HD, Krishnan I, Maroni G, Meyerovitz CV, Kerwin CM, Choi S, Richards WG, De Rienzo A, Tenen DG, Bueno R, Levantini E, Pittet MJ and Klein AM. Single-cell transcriptomics of human and mouse lung cancers reveals conserved myeloid populations across individuals and species. *Immunity* 2019; 50: 1317-1334, e1310.
- [25] Xie H, Yao J, Wang Y and Ni B. Exosome-transmitted circVMP1 facilitates the progression and cisplatin resistance of non-small cell lung cancer by targeting miR-524-5p-METTL3/SOX2 axis. *Drug Deliv* 2022; 29: 1257-1271.
- [26] Kustermann C, Narbute K, Movcana V, Parfejevs V, Rumnieks F, Kaukis P, Priedols M, Mikilps-Mikgelbs R, Mihailova M, Andersone S, Dzalts A, Bajo-Santos C, Krams A and Abols A. iPSC-derived lung and lung cancer organoid model to evaluate cisplatin encapsulated autologous iPSC-derived mesenchymal stromal cell-isolated extracellular vesicles. *Stem Cell Res Ther* 2024; 15: 246.
- [27] Zhou Z, Qu C, Zhou P, Zhou Q, Li D, Wu X and Yang L. Extracellular vesicles activated cancer-associated fibroblasts promote lung cancer metastasis through mitophagy and mtDNA transfer. *J Exp Clin Cancer Res* 2024; 43: 158.
- [28] Zhang Y, Hu K, Qu Z, Xie Z and Tian F. ADAMTS8 inhibited lung cancer progression through suppressing VEGFA. *Biochem Biophys Res Commun* 2022; 598: 1-8.
- [29] Li L, Yang L, Fan Z, Xue W, Shen Z, Yuan Y, Sun X, Wang D, Lian J, Wang L, Zhao J and Zhang Y. Hypoxia-induced GBE1 expression promotes tumor progression through metabolic reprogramming in lung adenocarcinoma. *Signal Transduct Target Ther* 2020; 5: 54.
- [30] Yang X, Zhang Q, Wei L and Liu K. HIF1A/PCDH7 axis mediates fatty acid synthesis and metabolism to inhibit lung adenocarcinoma anoikis. *J Biochem Mol Toxicol* 2024; 38: e70001.
- [31] Geng P, Ye F, Dou P, Hu C, He J, Zhao J, Li Q, Bao M, Li X, Liu X and Xu G. HIF-1 α -HPRT1 axis promotes tumorigenesis and gefitinib resistance by enhancing purine metabolism in EGFR-mutant lung adenocarcinoma. *J Exp Clin Cancer Res* 2024; 43: 269.

- [32] Wang Y, Cheng S, Fleishman JS, Chen J, Tang H, Chen ZS, Chen W and Ding M. Targeting anoikis resistance as a strategy for cancer therapy. *Drug Resist Updat* 2024; 75: 101099.
- [33] Grossmann J. Molecular mechanisms of “detachment-induced apoptosis–Anoikis”. *Apoptosis* 2002; 7: 247-260.
- [34] Chunhacha P, Sriuranpong V and Chanvorachote P. Epithelial-mesenchymal transition mediates anoikis resistance and enhances invasion in pleural effusion-derived human lung cancer cells. *Oncol Lett* 2013; 5: 1043-1047.
- [35] Kang J, Chun J, Hwang JS, Pan C, Li J, Boese AC, Young I, Malin CM, Kang Y, Gibbons DL, Sica G, Fu H, Ramalingam SS, Jin L and Kang S. EGFR-phosphorylated GDH1 harmonizes with RSK2 to drive CREB activation and tumor metastasis in EGFR-activated lung cancer. *Cell Rep* 2022; 41: 111827.
- [36] Zhao L, Yu Z and Zhao B. Mechanism of VIPR1 gene regulating human lung adenocarcinoma H1299 cells. *Med Oncol* 2019; 36: 91.
- [37] Fu Y, Liu S, Rodrigues RM, Han Y, Guo C, Zhu Z, He Y, Mackowiak B, Feng D, Gao B, Zeng S and Shen H. Activation of VIPR1 suppresses hepatocellular carcinoma progression by regulating arginine and pyrimidine metabolism. *Int J Biol Sci* 2022; 18: 4341-4356.
- [38] Zhang K, Tian R, Wang G, Zhang J, Ma H, Hu X, Xi J and Wang G. ADAMTS8 inhibits cell proliferation and invasion, and induces apoptosis in breast cancer. *Onco Targets Ther* 2020; 13: 8373-8382.
- [39] He Q, Qu M, Shen T, Xu Y, Luo J, Tan D, Xu C, Barkat MQ, Zeng LH and Wu X. Suppression of VEGFD expression by S-nitrosylation promotes the development of lung adenocarcinoma. *J Exp Clin Cancer Res* 2022; 41: 239.
- [40] Cheon I, Lee S, Oh S and Ahn YH. miR-200-mediated inactivation of cancer-associated fibroblasts via targeting of NRP2-VEGFR signaling attenuates lung cancer invasion and metastasis. *Mol Ther Nucleic Acids* 2024; 35: 102194.
- [41] Qiao H, Feng Y and Tang H. COL6A6 inhibits the proliferation and metastasis of non-small cell lung cancer through the JAK signalling pathway. *Transl Cancer Res* 2021; 10: 4514-4522.
- [42] Lu X, Shen J, Huang S, Wang H and Liu D. Down-regulation of CLEC3B facilitates epithelial-mesenchymal transition, migration and invasion of lung adenocarcinoma cells. *Tissue Cell* 2022; 76: 101802.
- [43] Seyed Abutorabi E, Irani S, Yaghmaie M and Ghaffari SH. Abemaciclib (CDK4/6 Inhibitor) blockade induces cytotoxicity in human anaplastic thyroid carcinoma cells. *Rep Biochem Mol Biol* 2020; 8: 438-445.
- [44] Eguchi R, Fujita Y, Tabata C, Ogawa H, Wakabayashi I, Nakano T and Fujimori Y. Inhibition of Src family kinases overcomes anoikis resistance induced by spheroid formation and facilitates cisplatin-induced apoptosis in human mesothelioma cells. *Oncol Rep* 2015; 34: 2305-2310.
- [45] Lee H, Jang Y, Park S, Jang H, Park EJ, Kim HJ and Kim H. Development and evaluation of a CEACAM6-targeting theranostic nanomedicine for photoacoustic-based diagnosis and chemotherapy of metastatic cancer. *Theranostics* 2018; 8: 4247-4261.

Machine learning identifies hypoxia-anoikis genes for LUAD prognosis



Supplementary Figure 1. IC50 of chemotherapy drugs and targeted drugs. A. Heatmap of correlations between prognostic genes, risk scores, and chemotherapy drug sensitivity. B. Heatmap of correlations between prognostic genes, risk scores, and targeted drug sensitivity. * $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$.