



Prognostic role of stress granule-related gene signatures in pancreatic ductal adenocarcinoma: insights from 101-combination machine learning and single-cell sequencing

Lei Wang, PhD^a, Linhao Zong^{b,*}, Yuanyuan Wang, MD^a, Yanting Jiang, MMed^a, Miao Guan, PhD^{b,*}

Background: Pancreatic ductal adenocarcinoma (PDAC) is an exceptionally aggressive malignancy of the digestive system, characterized by a fibrotic microenvironment that serves as an ideal model for studying stress granules (SGs). This study aimed to investigate SG-related mechanisms in PDAC, with particular focus on risk stratification and therapeutic strategies.

Methods: PDAC-related datasets were retrieved from The Cancer Genome Atlas and Gene Expression Omnibus databases. Differential expression analysis, univariate Cox regression, and 101 algorithmic combinations from 10 machine learning methods were employed to identify prognostic SG-related genes (SGRGs) and construct a risk model. Prognostic analyses were further extended through independent prognostic evaluation, nomogram development, immune microenvironment profiling, drug sensitivity testing, and enrichment analysis. Additionally, GSE197177 was examined to identify key cell types and perform pseudo-time and cell communication analyses.

Results: A risk model based on four prognostic SGRGs (LAMA3, ITGA6, COL17A1, and TOP2A) was developed, demonstrating superior predictive accuracy for PDAC prognosis. A nomogram incorporating age, N stage, and risk score was constructed, showing robust prognostic capacity. Further analyses revealed that immune cells, such as M0 macrophages and CD8 T cells, as well as drug sensitivities to ERK inhibitors and trametinib, were associated with risk stratification in PDAC patients. ITGA6 was notably enriched in the “regulation of glycolytic process” pathway. Pseudo-time analysis indicated a significant correlation between the expression of prognostic SGRGs and the differentiation status of key ductal cells, while cell communication analysis highlighted strong interactions between ductal cells and fibroblasts.

Conclusion: This study highlights the pivotal role of SGs in PDAC progression. A novel prognostic signature based on SGRGs was developed and validated, offering substantial potential for predicting patient outcomes in PDAC.

Keywords: machine learning, pancreatic ductal adenocarcinoma, prognostic signature, single-cell RNA sequencing, stress granules

Introduction

Pancreatic ductal adenocarcinoma (PDAC), the dominant histological subtype of pancreatic cancer, constitutes more than 90% of cases^[1]. Known for its high malignancy and rapid progression, PDAC is associated with a dismal 5-year survival rate of less than 10%^[2]. Even after radical surgery, the recurrence rate remains

alarmingly high at 70–80%, leading to a poor overall prognosis^[3]. Current treatment options for PDAC include surgical resection, adjuvant chemotherapy, radiotherapy, and emerging immunotherapies^[4]. However, their overall effectiveness is limited by challenges such as delayed diagnosis, the lack of effective biomarkers for early screening, complex mechanisms of chemotherapy resistance, and slow advances in targeted drug development^[5]. Furthermore, the tumor's highly fibrotic microenvironment, marked by a dense extracellular matrix, hinders drug delivery, further diminishing the efficacy of existing treatments^[6,7]. Thus, the identification of reliable prognostic biomarkers and therapeutic targets and the development of personalized intervention strategies are critical to addressing the therapeutic challenges of PDAC and improving patient outcomes.

Stress granules (SGs) are cytoplasmic, non-membranous aggregates that form in response to various cellular stresses. They serve as dynamic centers for mRNA storage and protein synthesis regulation, facilitating cellular adaptation and

^aDepartment of Radiation Oncology, Lianyungang Clinical College of Nanjing Medical University (The First People's Hospital of Lianyungang), Lianyungang, Jiangsu, China and ^bJiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, Nanjing, Jiangsu, China

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

*Corresponding authors. Address: Jiangsu Key Laboratory for Biodiversity and Biotechnology, College of Life Sciences, Nanjing Normal University, 1 Wenyuan Rd, Nanjing, Jiangsu 210023, China. Tel.: +8618362935394. E-mail: 08326@njnu.edu.cn (M. Guan); Tel.: +8618806278729. E-mail: lichthao@163.com (L. Zong).

Copyright © 2026 The Author(s). Published by Wolters Kluwer Health, Inc. This is an open access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

International Journal of Surgery (2026) 112:7811–7831

Received 3 July 2025; Accepted 25 November 2025

Supplemental Digital Content is available for this article. Direct URL citations are provided in the HTML and PDF versions of this article on the journal's website, www.ijso.com/international-journal-of-surgery.

Published online 20 January 2026

<http://dx.doi.org/10.1097/JS9.0000000000004458>

survival^[8,9]. The fibrotic tumor microenvironment and KRAS mutation-driven metabolic dysregulation in PDAC make it an ideal model for SG research^[10]. For example, Hao *et al* demonstrated that increased BZW1 expression in PDAC cells promotes eIF2 α phosphorylation *via* the PERK-eIF2 α pathway. This phosphorylation inhibits translation initiation complex assembly, leading to the accumulation of mRNA-protein complexes and the initiation of SG formation. Thus, SG assembly significantly enhances PDAC cell survival under hypoxic and glucose-deprived conditions. Notably, high BZW1 expression is associated with poor prognosis in PDAC patients^[11]. Despite these findings, the prognostic value of SGs in PDAC remains poorly understood, and the underlying mechanisms driving their roles in PDAC progression and drug resistance are yet to be fully elucidated.

Single-cell RNA sequencing (scRNA-seq) is a high-throughput technology that enables precise characterization of cellular heterogeneity, elucidates intercellular communication patterns, and provides high-resolution dynamic transcriptomic profiling^[12,13]. In PDAC research, scRNA-seq allows for the analysis of gene expression changes across diverse cell types within the tumor microenvironment, including cancer cells, fibroblasts, and immune cells. This technique facilitates the identification of key cellular subpopulations and core regulatory pathways driving PDAC initiation and progression^[14]. Clinically, scRNA-seq contributes to patient stratification by identifying SG-related genes (SGRGs) linked to prognosis, offering a foundation for more accurate monitoring of drug responses and disease progression.

Machine learning (ML) algorithms are advanced computational methods that construct data-driven models for pattern recognition, prediction, and decision making^[15]. In recent years, integrated ML frameworks based on multialgorithm ensembles have made significant strides in PDAC research, providing a new paradigm for understanding PDAC heterogeneity and enabling personalized diagnosis and treatment^[16,17].

This study aimed to address existing knowledge gaps by developing a novel risk model based on SGRGs using an ensemble ML approach. By integrating comprehensive RNA-seq data from The Cancer Genome Atlas (TCGA), a risk model associated with patient survival was identified through robust ML algorithms. The prognostic utility of this model was validated in an external cohort from Gene Expression Omnibus (GEO), followed by a series of bioinformatics analyses, including immune microenvironment profiling, drug sensitivity testing, and functional enrichment analysis. Furthermore, scRNA-seq was employed to characterize the tumor microenvironment, assess cell-type-specific risk patterns, and explore intercellular communication networks, providing novel therapeutic targets and insights for PDAC management. All procedures conducted in this study were in line with REMARK criteria^[18]. In compliance with the TITAN Guidelines 2025, this research was carried out without the utilization of artificial intelligence (AI) tools^[19].

Materials and methods

Data acquisition

Transcriptome data from pancreatic adenocarcinoma (PAAD) patients were retrieved from the TCGA database (<http://cancer.genome.nih.gov/>, accessed 20 March 2025). The PDAC training set (TCGA-PAAD) consisted of 179 tumor tissue samples

HIGHLIGHTS

- Key prognostic SGRGs in PDAC (LAMA3, ITGA6, COL17A1, and TOP2A) were identified.
- The differences in immune cell composition between high- and low-risk groups were analyzed.
- Sensitive drugs specific to high- and low-risk groups were identified.
- Biological mechanisms underlying prognostic SGRGs in PDAC were uncovered.
- Correlation between prognostic SGRGs and differentiation status of ductal cells was discovered.

classified as the PDAC group and 4 adjacent normal tissue samples as the normal group. Among these, 178 tumor tissue samples had available survival data. Transcriptome data from the validation set GSE28735 and the scRNA-seq dataset GSE197177, both related to PDAC, were obtained from the GEO database (<http://www.ncbi.nlm.nih.gov/geo/>). GSE28735 retained 45 PDAC tumor tissue samples (42 with survival data) and 45 adjacent normal tissue samples (platform: GPL6244). GSE197177 (platform: GPL18573) included three PDAC tumor tissue samples and one adjacent normal tissue sample. Additionally, the GSE183795 dataset (GPL6244) (139 PDAC cancer tissue samples) and the GSE62452 dataset (GPL6244) (69 PDAC cancer tissue samples) were utilized for transcriptome validation. GSE212966 (GPL24676) (six PDAC cancer tissue samples and six adjacent normal tissue samples) served as the second scRNA-seq dataset. Moreover, 844 SGRGs were retrieved from the GeneCards database (<https://www.genecards.org/>) (Supplemental Digital Content Table S1, available at: <http://links.lww.com/JS9/G305>).

Determination and functional enrichment analysis of differentially expressed stress granule-related genes (DE-SGRGs)

Differentially expressed genes (DEGs) between PDAC and normal groups in the PDAC training set were determined using the DESeq2 package (v 1.40.2) ($|\log_2$ fold change (FC)| > 1 and $P < 0.05$)^[20]. The results were visualized using the ggplot2 (v 3.5.1) and ComplexHeatmap (v 2.16.0) packages to generate volcano plots and heatmaps, respectively^[21,22]. To assess the consistency of the findings, a random sampling approach was applied with a random seed (seed = 5772), where 4 cases were randomly selected from the tumor samples and compared with TCGA normal samples ($n = 4$). This process was repeated five times. The overlap rate for each sampling (significant DEGs from random sampling/significant DEGs without sampling) was calculated and visualized as a bar chart using ggplot2 (v 3.5.1). The intersection of DEGs and SGRGs was identified to obtain DE-SGRGs using the ggvenn package (v 0.1.10)^[23]. To explore the biological functions and processes of DE-SGRGs in PDAC, Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses were conducted using the clusterProfiler package (v 4.15.0.003) (adj. $P < 0.05$)^[24]. The background gene set was set to SGRGs, and GO enrichment results were displayed (adj. $P < 0.05$). Finally, a protein-protein interaction (PPI) network was constructed using the STRING

database (<http://www.string-db.org/>) (interaction score = 0.7) to examine the interaction relationships among proteins associated with DE-SGRGs.

Integrated machine learning algorithm

In cancer tissue samples from the PDAC-training set, univariate Cox regression analysis [hazard ratio (HR) $\neq 1$, $P < 0.001$] and the proportional hazards (PH) assumption test ($P > 0.05$) were conducted on the DE-SGRGs to select candidate prognostic SGRGs (SGRGs1) using the survival package (v 3.7-0)^[25]. Due to the small sample size, traditional ML algorithms were selected. To enhance precision and stability, 10 ML algorithms and 101 algorithmic combinations were employed. Model performance was rigorously evaluated using 10-fold cross-validation. The ML.Dev.Prog.Sig function of the Mime1 package (v 0.0.0.9000) (mode = all) was utilized to integrate combinations^[26,27]. The algorithms included LASSO regression, Ridge regression stepwise, COX regression, elastic net, RSF, CoxBoost, PLS-COX, supervised PCA, GBRM, and survival SVM. For each model, the concordance index (C-index) was computed for both the PDAC-training set and GSE28735. For Lasso/Ridge regression, the optimal regularization parameter lambda was automatically selected *via* 10-fold cross-validation in glmnet. In the random forest model, the minimum number of samples in terminal nodes was set to 5 (nodesize = 5). For the GBM model, early stopping was applied by monitoring cross-validation errors, with the iteration number yielding the smallest error selected. To prevent overfitting, the following control measures were implemented: (1) screening *via* univariate Cox regression ($P < 0.05$) combined with multivariate stepwise regression; (2) regularization (L1/L2 penalty); (3) controlling model complexity; and (4) evaluation using independent validation sets (GSE183795 and GSE62452). The model with the highest average C-index was deemed optimal and used for screening candidate prognostic SGRGs (SGRGs2). The ggvenn package (v 0.1.10) was employed to identify the intersection of the candidate prognostic SGRGs, yielding the final set of prognostic SGRGs.

Construction and verification of prognostic SGRG signature

The prognostic SGRGs were then used to construct the risk model, defined by the following formula:

$$\text{risk score} = \sum_{i=1}^N \text{coef}(\text{gene}_i) \times \text{expr}(\text{gene}_i)$$

where coef represents the coefficient, and expr denotes the expression value of each gene. Subsequently, cancer tissue samples with survival data from the PDAC-training set, GSE28735, GSE183795, and GSE62452 were classified into the high-risk group (HRG, risk score \geq optimal cutoff value) and the low-risk group (LRG, risk score $<$ optimal cutoff value) groups. The tynarray package (v 2.4.2) was used to visualize the distribution of risk scores, patient survival conditions, and the expression levels of prognostic SGRGs between the two risk groups^[28]. Kaplan–Meier (KM) and receiver operating characteristic (ROC) curves were generated using the survminer package (v 0.4.9) and survivalROC package (v 1.0.3.1), respectively, to assess the prognostic performance of the risk model^[29,30]. If the area under the curve (AUC) values for 1-, 2-, and 3-year survival

exceeded 0.6, the model was considered to exhibit moderate discriminatory power. Additionally, ROC curves for individual prognostic SGRGs and the risk model were plotted at 1-, 2-, and 3-year time points to evaluate the timeliness and stability of the model's predictions.

Analyses of the clinical characteristics and nomogram

To evaluate the prognostic significance of the risk model for overall survival (OS) in PDAC-training set samples with available survival data, univariate and multivariate Cox regression analyses (HR $\neq 1$, $P < 0.05$) with PH assumption tests ($P > 0.05$) were performed using the survival package (v 3.7-0). These analyses incorporated clinical characteristics and risk scores to identify independent prognostic factors.

Subsequently, a nomogram was developed using the rms package (v 6.8-1) to predict 1-, 2-, and 3-year survival in PDAC patients^[31]. The nomogram assigned points to each factor, with each factor corresponding to a specific number of points. The total points from all factors were summed, allowing for the prediction of the 1-, 2-, and 3-year survival probabilities. To assess the reliability of the nomogram, calibration curves were generated using the rms package (v 6.8-1).

Immune microenvironment analysis

The progression and drug resistance of PDAC are strongly linked to the high heterogeneity of immune cells within the tumor microenvironment^[14]. The CIBERSORT algorithm was used to calculate the infiltration abundance of 22 immune cell types in HRG and LRG of the PDAC-training set (excluding samples with $P > 0.05$). Differential immune cells (DICs) were identified by performing a Wilcoxon test ($P < 0.05$)^[32]. Spearman correlation analyses between DICs and prognostic SGRGs were conducted using the psych package (v 2.4.6) (|correlation coefficient (cor)| > 0.30 and $P < 0.05$)^[33]. The ESTIMATE package (v 1.0.13) was used to compute ESTIMATE scores, immune scores, and stromal scores for the two risk groups, with statistical significance set at $P < 0.05$ ^[34]. Immune checkpoints, which suppress immune cell activity against tumor cells, were assessed by performing a Wilcoxon test ($P < 0.05$)^[35].

Drug sensitivity analyses

For drug sensitivity analysis, the GDSC database (<https://www.cancerrxgene.org/>) was used to identify therapeutic recommendations for PDAC management. The oncoPredict package (v 1.2) was employed to calculate the half-maximal inhibitory concentration (IC₅₀) of 198 common chemotherapeutic and molecular-targeted agents in the PDAC-training set, which was used to infer drug sensitivity^[36]. The Wilcoxon test was applied to compare the variations in drug sensitivity between HRG and LRG ($P < 0.05$). The top five drugs exhibiting the greatest significant differences in sensitivity were selected and visualized in a box plot.

Mutation status analysis

Somatic mutation analysis plays a pivotal role in identifying driver mutations and therapy-related mutations in tumor cells. This analysis provides valuable insights for the diagnosis, prognosis, and development of personalized treatment strategies for

PDAC^[37]. To investigate somatic mutation variations between HRG and LRG in the PDAC-training set, the plotmafSummary function from the maftools package (v 2.16.0) was employed to analyze mutation information, including mutation count, distribution of mutation types, and highly mutated genes^[38].

Gene set enrichment analysis (GSEA)

Gene set enrichment analysis (GSEA) was performed to further explore the potential biological functions of prognostic SGRGs in cancer tissue samples from the PDAC-training set with available survival data. The Spearman correlation values for each prognostic SGRG with all other genes were computed using the psych package (v 2.4.6). These correlations were ranked in descending order, and the corresponding genome for each prognostic SGRG was identified. The reference gene set “c2.cp.all.v2022.1.Hs.symbols.gmt” from the Molecular Signature Database (MSigDB) (<https://www.gsea-msigdb.org/gsea/msigdb/>) was used, and GSEA was performed with the clusterProfiler package (v 4.15.0.003) ($P < 0.05$).

The scRNA-seq data processing

For the scRNA-seq data from all samples in GSE197177, data were filtered using the Read10X and CreateSeuratObject functions from the Seurat package (v 5.1.0) to select high-quality cells (nFeature RNA between 200 and 10 000, nCount RNA > 1000, and mitochondrial proportion < 20%)^[39]. The data were then standardized using the LogNormalize function. The FindVariableFeatures function was applied to extract and display the top 3000 highly variable genes (HVGs). Principal component analysis (PCA) was performed using the ScaleData, JackStraw, and ElbowPlot functions to determine the top principal components (PCs) ($P < 0.05$). Uniform Manifold Approximation and Projection (UMAP) was used for dimensionality reduction and cell clustering through the RunUMAP function (resolution = 0.5). Cells were annotated according to marker genes using the FindAllMarkers function, PDAC-related single-cell literature, and the CellMarker2.0 database (<http://117.50.127.228/CellMarker/index.html>)^[40]. The expression of marker genes across different cell types was visualized using a bubble plot. The proportions of annotated cells in each sample were presented, and cells with differential expression of each prognostic SGRG between PDAC and normal groups were selected as key cells ($P < 0.05$). Similarly, the same analysis was applied to the GSE212966 single-cell dataset. The scRNA-seq data were filtered using the Seurat package (v 5.1.0) to select high-quality cells (nFeature RNA between 200 and 7500, nCount RNA between 1000 and 80 000, and mitochondrial proportion < 20%). The remaining analysis steps were consistent with those performed on GSE197177.

Pseudo-time and cell communication analysis

In all samples from GSE197177, key cells underwent secondary dimensionality reduction and clustering as described in section “The scRNA-seq data processing,” aiming to explore the dynamics and temporal trajectories of prognostic SGRG expression across cell subtypes and to infer cell evolution during PDAC progression. The key cells were classified into distinct subgroups, and pseudo-time analysis was performed using the monocle package (v 2.28.0) to investigate their differentiation

states and subtypes^[41]. The variations in prognostic SGRG expression levels were evaluated during key cell differentiation. To assess communication between the key cell clusters and other cell populations in PDAC and normal groups, the CellChat package (v 1.6.0) was applied^[42].

Expression levels of prognostic SGRGs

Differences in the expression levels of prognostic SGRGs between PDAC and normal groups were analyzed in the PDAC-training set and GSE28735 datasets using the Wilcoxon test ($P < 0.05$).

Results

Identification and functional assessment of 90 DE-SGRGs

The PDAC-training set identified 3456 DEGs, consisting of 1860 upregulated genes and 1596 downregulated genes in the PDAC group. The volcano plot highlighted the top 10 up- and down-regulated DEGs (Fig. 1A). A heatmap revealed diverse expression patterns between the PDAC and normal groups (Fig. 1B). The overlap rates of these five samplings ranged from 40.7 to 52.1%, showing significant differences from the DEGs without sampling (Supplemental Digital Content Figure S1, available at: <http://links.lww.com/JS9/G304>). The intersection of the 3456 DEGs and 844 SGRGs revealed 90 DE-SGRGs (Fig. 1C, Supplemental Digital Content Table S2, available at: <http://links.lww.com/JS9/G305>). Subsequent GO and KEGG analyses of the 90 DE-SGRGs elucidated their molecular biological processes. A total of 1150 GO terms were enriched, with “reactive oxygen species metabolic process,” “leukocyte migration,” and “response to oxidative stress” enriched in 1024 biological process (BP) entries. In the cellular component (CC) category, 69 entries, including “endocytic vesicle” and “plasma membrane raft,” were identified. The molecular function (MF) category contained 57 entries, indicating a strong association of DE-SGRGs with “amyloid-beta binding” and “protein-lipid complex binding” (adj. $P < 0.05$, Fig. 1D, Supplemental Digital Content Table S3, available at: <http://links.lww.com/JS9/G305>). When the background gene set was switched to SGRGs, DE-SGRGs were enriched in six GO terms, such as “plasma lipoprotein particle clearance,” “cell periphery,” and “transporter activity” (adj. $P < 0.05$, Fig. 1E, Supplemental Digital Content Table S4, available at: <http://links.lww.com/JS9/G305>). KEGG analysis revealed that DE-SGRGs were enriched in 34 KEGG pathways, including the “AGE-RAGE signaling pathway in diabetic complications,” “leukocyte transendothelial migration,” and “ECM-receptor interaction” (adj. $P < 0.05$, Fig. 1F, Supplemental Digital Content Table S5, available at: <http://links.lww.com/JS9/G305>). A PPI network was constructed to investigate the interactions among the 90 DE-SGRGs, resulting in a network consisting of 72 proteins and 280 interaction pairs. Within this network, APOE, ITGB2, REN, CD36, PIK3CG, and ICAM1 exhibited extensive interactions with other proteins (Fig. 1G). These findings offer valuable insights into the molecular mechanisms driving PDAC.

Construction of SGRG-related risk model based on an integrated machine learning program

Univariate Cox regression analysis and PH assumption tests identified 12 candidate prognostic SGRGs1 significantly associated

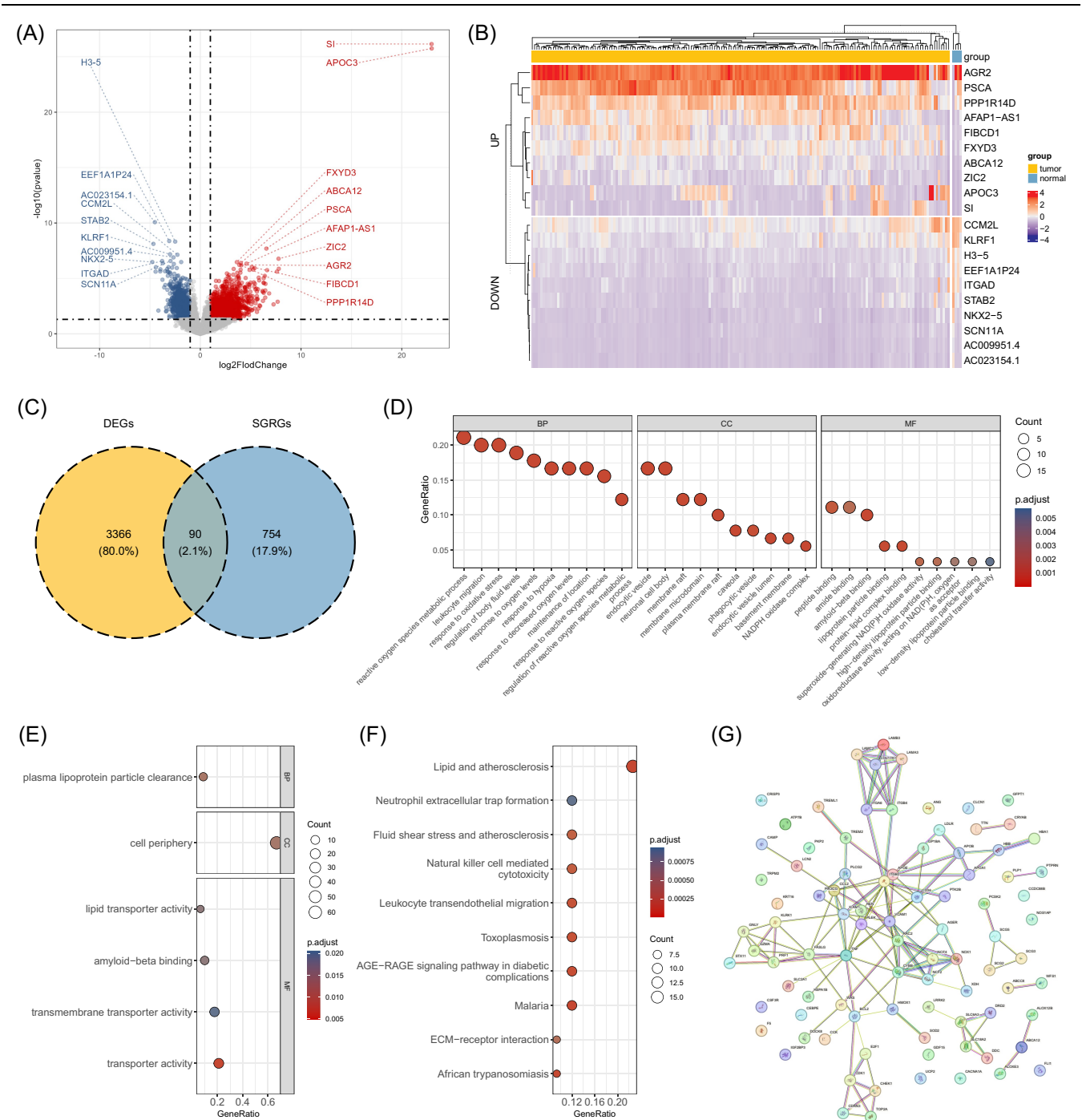


Figure 1. (A) Volcano plot of DEGs in the PDAC-training set. (B) Heatmap of DEGs in the PDAC-training set. (C) Venn plot of DEGs and SGRGs. (D and E) GO enrichment analysis of DE-SGRGs with the background gene set as (D) the universe gene set and (E) SGRGs. (F) KEGG enrichment analysis of DE-SGRGs with the background gene set as the universe gene set. (G) PPI network of DE-SGRGs.

with OS in PDAC patients from the PDAC training set. All candidate prognostic SGRGs1, including LAMA3, ITGA6, COL17A1, and TOP2A, were considered risk factors (HR > 1) (Fig. 2A, Supplemental Digital Content Figure S2, available at: <http://links.lww.com/JS9/G304>). In both the PDAC-training set and GSE28735, 101 prediction models were fitted, and the C-index for each model was computed. Notably, 14 models exhibited the highest C-index of 0.61 in the validation cohort

and a relatively high average C-index of 0.64 across all cohorts (Fig. 2B). Furthermore, four candidate prognostic SGRGs2 were identified: LAMA3, ITGA6, COL17A1, and TOP2A. The intersection of candidate prognostic SGRGs1 and SGRGs2 revealed four prognostic SGRGs (LAMA3, ITGA6, COL17A1, and TOP2A) (Fig. 2C). The formula for the risk model in PDAC patients was determined as: risk score = (0.228) × LAMA3 expression + (−0.399) × ITGA6 expression + (0.212) × COL17A1

expression + (0.485) × TOP2A expression. To evaluate the performance of this risk model, PDAC samples were categorized into high-risk (HRG) and low-risk (LRG) groups based on an optimal cutoff value for the risk score of 1.933407 (high/low risk patients = 78/100) in the PDAC-training set. The risk profile

and survival status plot showed that in HRG, as risk scores increased, survival time decreased, and the number of deaths increased correspondingly. The expression levels of LAMA3, ITGA6, COL17A1, and TOP2A were higher in HRG (Fig. 2D). KM curves revealed a lower survival probability in HRG,

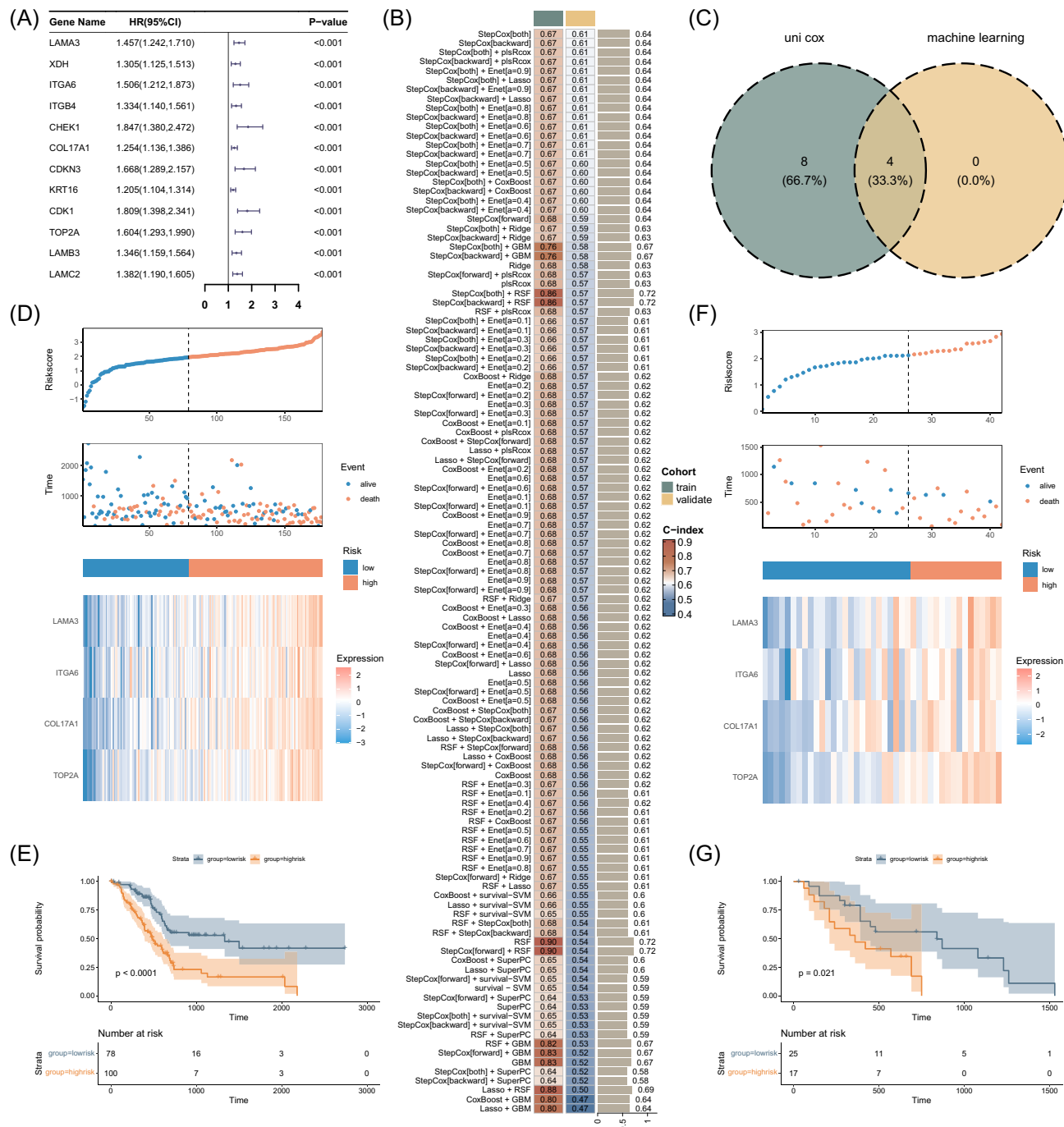


Figure 2. (A) Forest plot of univariate Cox regression analysis for candidate prognostic genes. (B) Heatmap of C-index results for machine learning model combinations. (C) Venn diagram showing the intersection of candidate gene sets selected by univariate Cox analysis and machine learning. (D and E) In the training set, (D) risk score curves for high/low-risk groups of PDAC (top); survival status of patients in high/low-risk groups (middle); heatmap of the expression of prognostic genes (bottom); (E) K-M curves evaluating the difference in overall survival (OS) between high/low-risk groups. (F and G) In the testing set, (F) risk score curves for high/low-risk groups of PDAC (top); survival status of patients in high/low-risk groups (middle); heatmap of the expression of prognostic genes (bottom); (G) K-M curves evaluating the difference in OS between high/low-risk groups.

indicating that PDAC patients with high risk had a poorer prognosis ($P < 0.0001$, Fig. 2E). This finding was further validated in the GSE28735 cohort (high/low risk patients = 25/17) ($P < 0.05$) (Fig. 2F and G). These results suggest that the risk model can accurately differentiate the prognosis of patients in different risk groups, highlighting its potential as a clinical tool for prognostic stratification.

Confirmation of SGRG-related risk model

The effectiveness and sensitivity of the established risk model were further validated within the PDAC-training set using AUC values. The model demonstrated strong predictive accuracy for

OS in PDAC patients, with AUCs of 0.70, 0.73, and 0.75 for 1-, 2-, and 3-year prognoses, respectively (Fig. 3A). Similarly, the AUC values for 1-, 2-, and 3-year survival in GSE28735 were 0.64, 0.78, and 0.83, respectively, indicating that the risk model exhibited substantial effectiveness and robustness (Fig. 3B). The diagnostic performance of prognostic SGRGs was assessed using ROC curves. The AUC values for multiple prognostic SGRGs at 1-, 2-, and 3-year time points were 0.698, 0.731, and 0.746, respectively, outperforming those of individual SGRGs. This suggests that the combined evaluation of multiple prognostic SGRGs may offer significant clinical value (Fig. 3C and D). Additionally, the risk model was validated in the GSE183795 dataset, yielding favorable results (Supplemental Digital

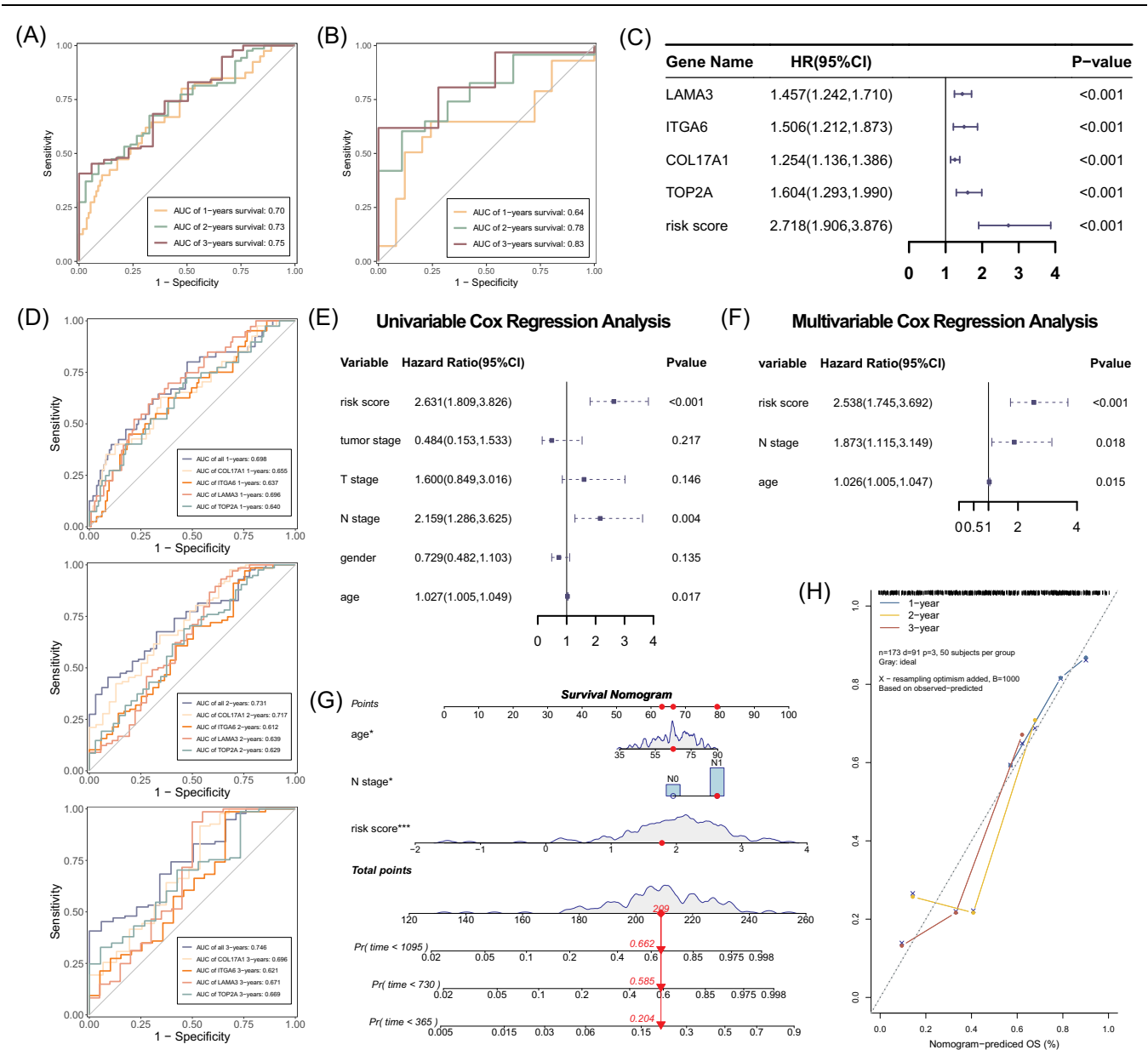


Figure 3. (A and B) ROC curves for 1-, 2-, and 3-year survival in (A) the training set and (B) the testing set. (C) Forest plot of single-gene risk scores. (D) ROC curves for individual genes in the training set. (E and F) Forest plots of (E) univariate and (F) multivariate Cox regression analyses for risk score and other factors. (G) Nomogram and (H) calibration curve of risk score, N stage, and age.

Content Figure S3, available at: <http://links.lww.com/JS9/G304>; however, its performance in the GSE62452 dataset was slightly weaker, with an AUC of 0.57 for 1-year survival (Supplemental Digital Content Figure S4, available at: <http://links.lww.com/JS9/G304>).

Independent prognostic analyses are crucial for developing robust clinical decision support systems. In the PDAC-training set, risk score, N stage, and age were identified as independent prognostic factors through univariate and multivariate Cox regression analyses and PH assumption tests (Fig. 3E and F and Supplemental Digital Content Figure S5A–B, available at: <http://links.lww.com/JS9/G304>). A nomogram was then developed incorporating risk score, N stage, and age to illustrate predictive accuracy and clinical utility for PDAC patients at 1-, 2-, and 3-year time points (Fig. 3G). Higher scores corresponded to a poorer prognosis for PDAC patients. Notably, calibration curves showed strong consistency between the survival probabilities predicted by the nomogram and the actual observed outcomes, demonstrating the model's excellent predictive precision (Fig. 3H). In conclusion, the SGRG-related risk model effectively forecasts PDAC prognosis and holds significant potential in supporting clinical decision making.

Infiltration landscape associated with risk scores in PDAC

Tumor-infiltrating immune cells significantly impact PDAC progression and are closely linked to clinical outcomes in PDAC patients. The abundance of immune cells was compared between HRG and LRG in the PDAC-training set (Fig. 4A). Among these immune cells, seven DICs showed significant differences ($P < 0.05$). HRG exhibited a notable increase in the infiltration of memory M0 macrophages and eosinophils, while the infiltration rates of CD8 T cells and monocytes were significantly lower in HRG ($P < 0.05$, Fig. 4B). A correlation heatmap revealed that M0 macrophages had the strongest positive correlation with COL17A1 ($\text{cor} = 0.33$, $P < 0.001$), whereas CD8 T cells showed the strongest negative correlation with ITGA6 ($\text{cor} = -0.30$, $P < 0.001$) (Fig. 4C). These results suggest that abnormal immune infiltration in PDAC, associated with SGs, offers valuable insights with critical clinical relevance.

The ESTIMATE algorithm results indicated that HRG had lower ESTIMATE, immune, and stromal scores ($P < 0.05$, Fig. 4D). Additionally, 13 immune checkpoint genes exhibited significant differences between the two risk groups ($P < 0.05$). For instance, nine immune checkpoint genes, including LGALS9, TNFSF9, HHLA2, and CD70, were significantly upregulated in HRG, while the expression levels of four immune checkpoint genes, including CD200, BTLA, and TNFSF14, were significantly lower in HRG ($P < 0.05$, Fig. 4E). These results suggest that patients in HRG may experience more pronounced immune evasion, potentially leading to less favorable outcomes with immunotherapeutic approaches.

SGRG signature as a potential tool for predicting treatment response

Drug sensitivity analysis of 148 anticancer drugs identified statistically significant differences in sensitivity between risk groups (Supplemental Digital Content Table S6, available at: <http://links.lww.com/JS9/G305>), with the top five drugs showing the most significant p -values selected for further analysis. Notably, low-risk patients demonstrated significantly higher sensitivities

to drugs such as JAK1_8709 and sabutoclax ($P < 0.0001$), suggesting that these drugs may offer enhanced therapeutic responses in LRG (Fig. 5A). In contrast, high-risk patients exhibited lower IC_{50} values for ERK and VX.11e ($P < 0.0001$), indicating that PDAC patients at higher risk are more likely to show reduced resistance to chemotherapy (Acetalax was excluded due to hepatotoxicity concerns, and ERK was chosen for presentation instead) (Fig. 5B). These results suggest that targeting prognostic SGRGs with these drugs could improve PDAC treatment outcomes.

Comparison regarding the mutation status between HRG and LRG

The gradual accumulation of gene mutations is closely linked to the onset and progression of tumors^[43]. Analyzing somatic mutation data provides deeper insights into the factors contributing to diverse clinical outcomes. The mutation analysis revealed that missense variants were the most common genetic alterations in the PDAC-training set. The waterfall plot demonstrated that typical gene mutations were present in both HRG and LRG. In HRG, the top three mutant genes were KRAS, TP53, and SMAD4, with KRAS exhibiting a dominant mutation rate of 77% and TP53 at 71% (Fig. 6A). In contrast, the top three mutant genes in LRG were TP53, KRAS, and TTN, with TP53 and KRAS mutation rates of 37 and 33%, respectively (Fig. 6B). Further examination of mutation types and single-nucleotide variation (SNV) patterns across the PDAC-training set and different risk groups revealed distinct mutation signatures associated with risk levels. Compared to HRG, LRG exhibited fewer mutations, a lower proportion of single-nucleotide polymorphisms (SNPs), and a higher proportion of C-T mutations (Fig. 6C and D). In summary, integrating mutation status with risk scores enhanced the predictive capacity for the prognosis of patients with PDAC.

Biological mechanisms associated with prognostic SGRGs in PDAC

GSEA plays a pivotal role in identifying extensive and consistent changes in biological pathways. GSEA indicated that LAMA3 was significantly enriched in 2274 pathways, including “sulfation” and “triglyceride-rich lipoprotein particle remodeling” ($P < 0.05$, Fig. 7A, Supplemental Digital Content Table S7, available at: <http://links.lww.com/JS9/G305>). ITGA6 was significantly enriched in 2571 pathways, such as “regulation of fibrinolysis,” “regulation of glutamate secretion,” and “glucose import” ($P < 0.05$, Fig. 7B, Supplemental Digital Content Table S8, available at: <http://links.lww.com/JS9/G305>). COL17A1 was enriched in 1737 pathways, including “interleukin 27 mediated signaling pathway,” “regulation of odontogenesis,” and “blood vessel remodeling” ($P < 0.05$, Fig. 7C, Supplemental Digital Content Table S9, available at: <http://links.lww.com/JS9/G305>). TOP2A was found to be enriched in 1955 pathways, such as “deoxyribonucleoside triphosphate biosynthetic process,” “regulation of transforming growth factor beta activation,” and “heparin biosynthetic process” ($P < 0.05$, Fig. 7D, Supplemental Digital Content Table S10, available at: <http://links.lww.com/JS9/G305>). These results suggest that prognostic SGRGs may influence SG formation by regulating these functional pathways during the pathophysiological process of PDAC.

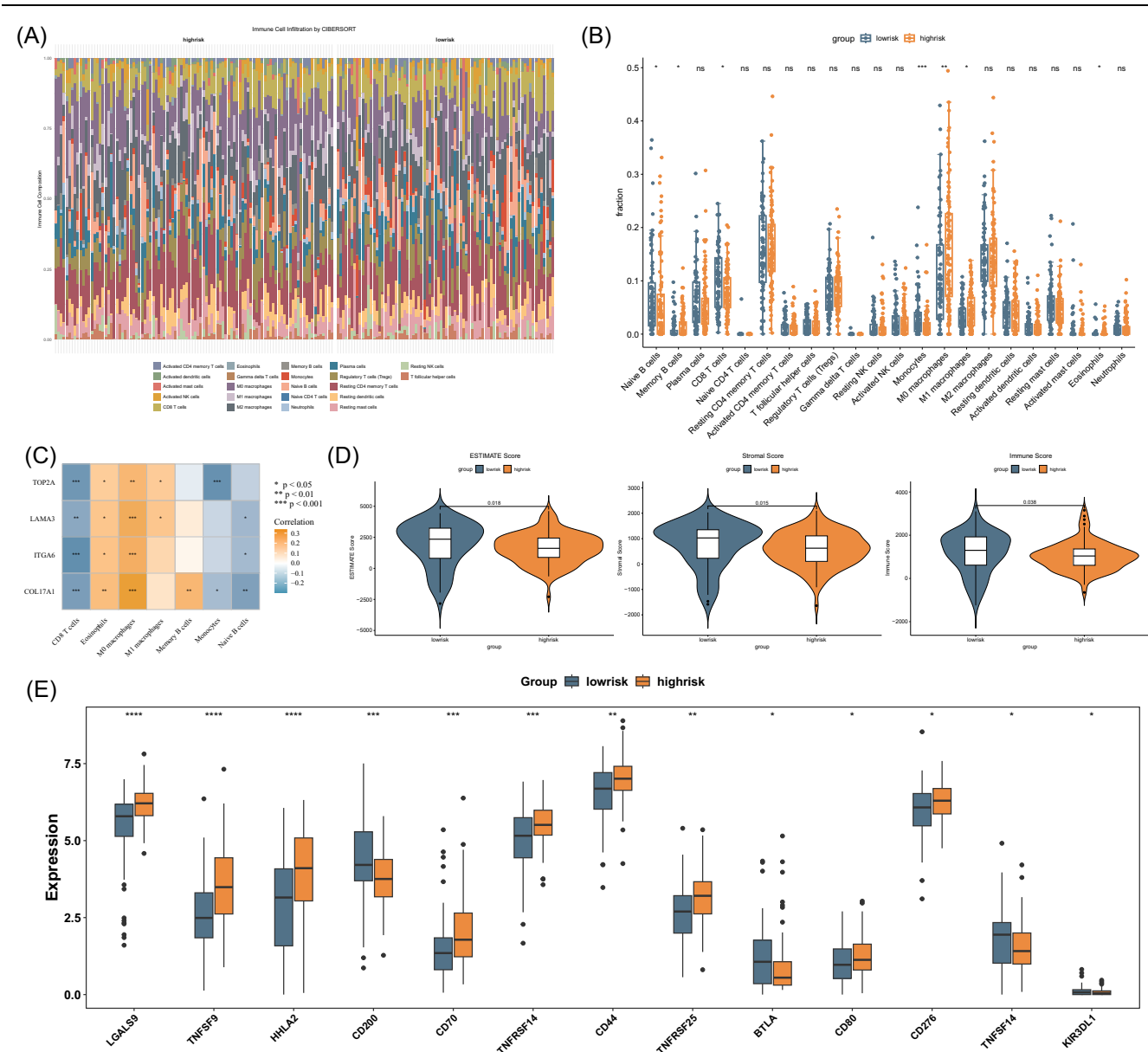


Figure 4. (A) Proportional plot of the infiltration abundance and (B) immune infiltration landscape of 22 immune cells in PDAC patients from high/low-risk groups. (C) Correlation plot among seven differentially infiltrated immune cells. (D) Comparison plot of ESTIMATE score, stromal score, and immune score from ESTIMATE analysis between high/low-risk groups. (E) Box plot showing the expression levels of classical immune checkpoint molecules in high/low-risk groups.

Comprehensive single-cell landscape of PDAC

Following integration and filtering of the original dataset, quality control (QC) was performed. Pre-QC data consisted of 49 331 cells and 33 538 genes, while post-QC data retained 32 004 cells and all 33 538 genes (Supplemental Digital Content Figure S6A, available at: <http://links.lww.com/JS9/G304>). The top 3000 HVGs were then selected for downstream analysis (Fig. 8A). PCA and the scree plot indicated that the optimal dimensionality was 20, so the top 20 PCs were retained for further analysis ($P < 0.05$, Supplemental Digital Content Figure S6B and C, available at: <http://links.lww.com/JS9/G304>). The cells were subsequently separated into 22 distinct clusters (Fig. 8B). These clusters were annotated into 11 cell types: ductal cells, mast cells, fibroblasts, macrophages, natural

killer (NK) cells, T cells, plasma cells, acinar cells, endothelial cells, B cells, and endocrine cells (Fig. 8C). A bubble plot was generated to visualize the expression of marker genes across these cell types (Fig. 8D, Supplemental Digital Content Table S11, available at: <http://links.lww.com/JS9/G305>). The histogram displayed the relative abundance of the 11 annotated cell types, with T cells and ductal cells showing the highest abundance in PDAC samples (Fig. 8E). Notably, the expression of prognostic SGRGs in ductal cells exhibited significant differences between the two groups ($P < 0.01$), leading to the selection of ductal cells as the key cell type for further analysis (Fig. 8F). In the GSE212966 dataset, following integration and filtering, QC was also performed. Pre-QC data included 71 074 cells and 36 601 genes, while post-QC data retained 53 091 cells and

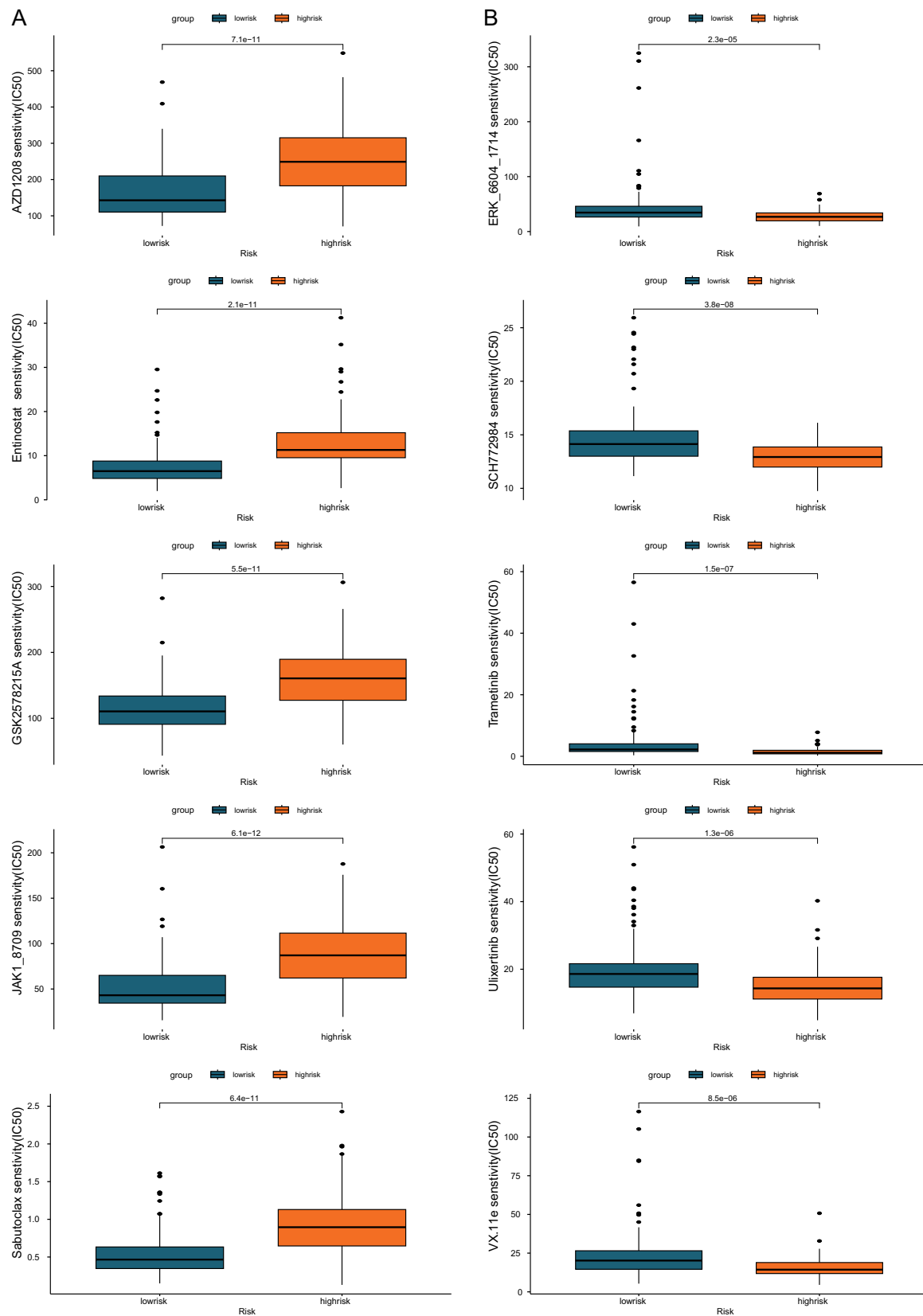


Figure 5. Sensitivity score plot of common antitumor drugs (top five drugs) for PDAC patients with (A) low-risk group or (B) high-risk group showing greater sensitivities.



Figure 6. (A and B) Waterfall plots of tumor mutation burden analysis for (A) the high-risk group and (B) the low-risk group. The top section displays the tumor mutation burden for each sample, representing the number of somatic mutations per megabase (per 1 million bases). The waterfall-like section in the middle illustrates the mutation status of each gene in each sample, with each row corresponding to a gene and each column to a sample. Different colors indicate distinct mutation types. The bar plot on the right shows the proportion of mutations containing mutations in the gene and the composition of mutation types. (C and D) Summary plots of tumor mutation burden analysis for (C) the high-risk group and (D) the low-risk group. From top to bottom and left to right, the plots include: a statistical chart of various mutation classifications (with mutation classifications from Ensembl on the y-axis and mutation counts on the x-axis), a statistical chart of mutation types (with mutation types on the x-axis, where SNP denotes point mutations, INS denotes insertions, and DEL denotes deletions, and mutation counts on the y-axis), a statistical chart of base changes (showing the types of SNP changes, e.g., T > G indicates a change from base T to G, with the percentage of each change type on the y-axis), a statistical chart of the number of mutations per sample (with mutation counts on the y-axis, sample numbers on the x-axis, and colors indicating mutation classifications as in the mutation classification statistical chart), a box plot of various mutation classifications across samples, the mutation types contained in the top 10 genes with the most mutations, and the sample proportions.

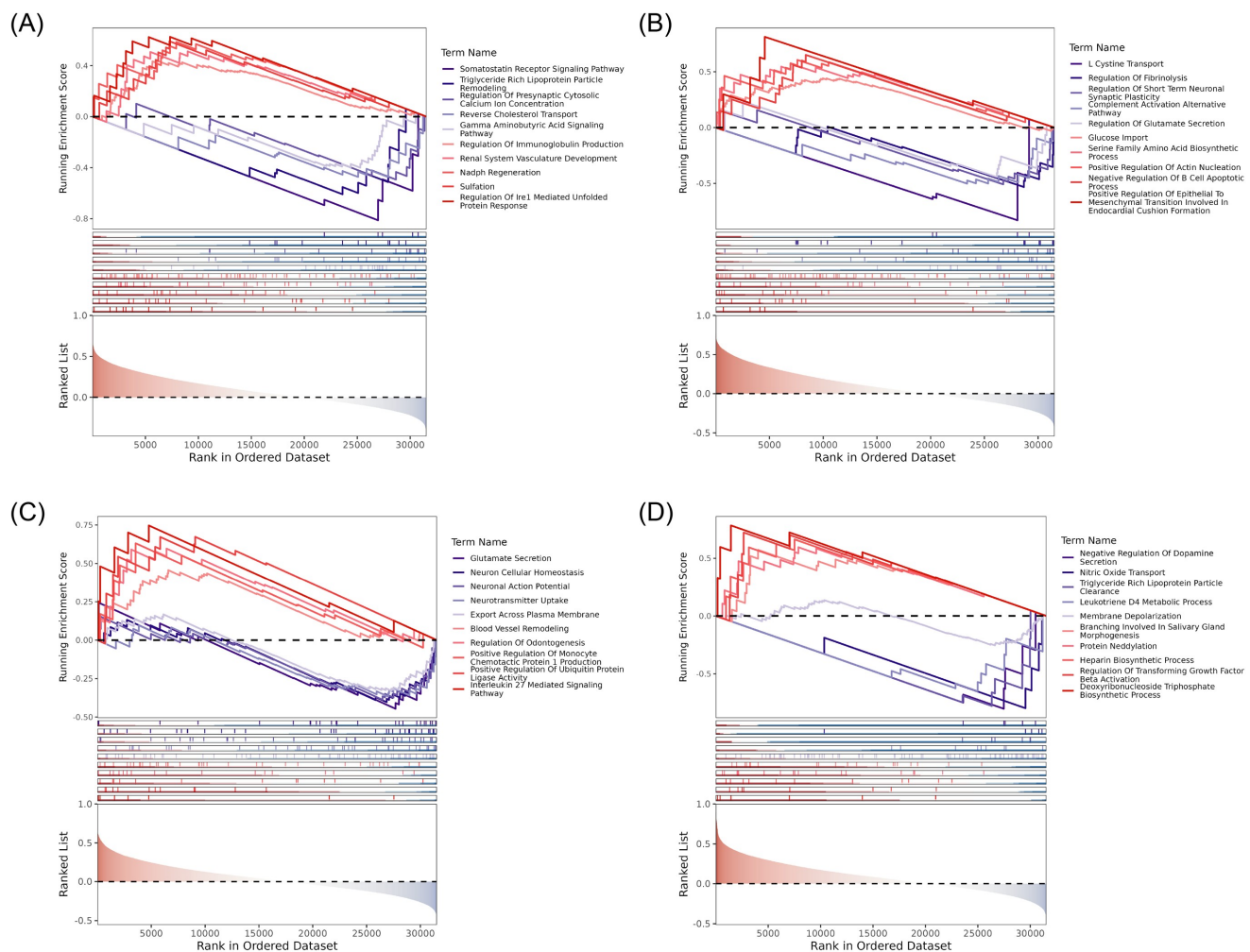


Figure 7. GSEA results of (A) LAMA3, (B) ITGA6, (C) COL17A1, and (D) TOP2A.

all 36 601 genes. Supplemental Digital Content Figure S7A, available at: <http://links.lww.com/JS9/G304>). The top 3000 HVGs were selected for downstream analysis (Fig. 9A). PCA and scree plots determined that the optimal dimensionality was 30, and the top 30 PCs were retained for subsequent analysis ($P < 0.05$, Supplemental Digital Content Figure S7B and C, available at: <http://links.lww.com/JS9/G304>). The cells were separated into 39 distinct clusters (Fig. 9B), which were annotated into the same 11 cell types as in the previous dataset (Fig. 9C). A bubble plot was generated to visualize marker gene expression (Fig. 9D). The histogram again showed that T cells and ductal cells were the most abundant cell types in PDAC samples (Fig. 9E). Given that prognostic SGRGs were highly expressed in ductal cells, these cells were selected as the key cell type for further investigation (Fig. 9F). The identification of ductal cells as key cells in both single-cell datasets demonstrated the robustness and stability of the analysis results.

Role of prognostic SGRGs and cell communication landscape in ductal cells

Following the identification of ductal cells as key cell types, pseudo-time analysis was performed to infer their differentiation

trajectories, revealing six distinct subtypes: RPS3+, TFF1+, MKI67+, CEACAM5+, CEACAM6+, and MALAT1 + ductal cells (Fig. 10A). Notably, CEACAM6+ and RPS3 + ductal cells differentiated earlier, whereas MALAT1+ and MKI67 + ductal cells underwent later differentiation. The differentiation process of ductal cells progressed through three distinct states, highlighting the inherent heterogeneity within ductal cell populations. Specifically, state 1 differentiated earlier, while state three exhibited later differentiation (Fig. 10B–D). Furthermore, the expression patterns of prognostic SGRGs across the pseudo-time trajectories were explored in ductal cells. As differentiation progressed, COL17A1 and ITGA6 expression levels increased, while LAMA3 expression exhibited an initial rise followed by a decline. TOP2A expression initially decreased, then nearly ceased, and later rebounded (Fig. 10E).

The cell communication network diagram illustrated the number and strength of interactions among the annotated cell types, demonstrating that ductal cells communicated with several other cell types. Notably, the number and intensity of cell–cell interactions in the PDAC group were significantly higher than those observed in the normal group (Fig. 11A). Strong interactions between ductal cells and

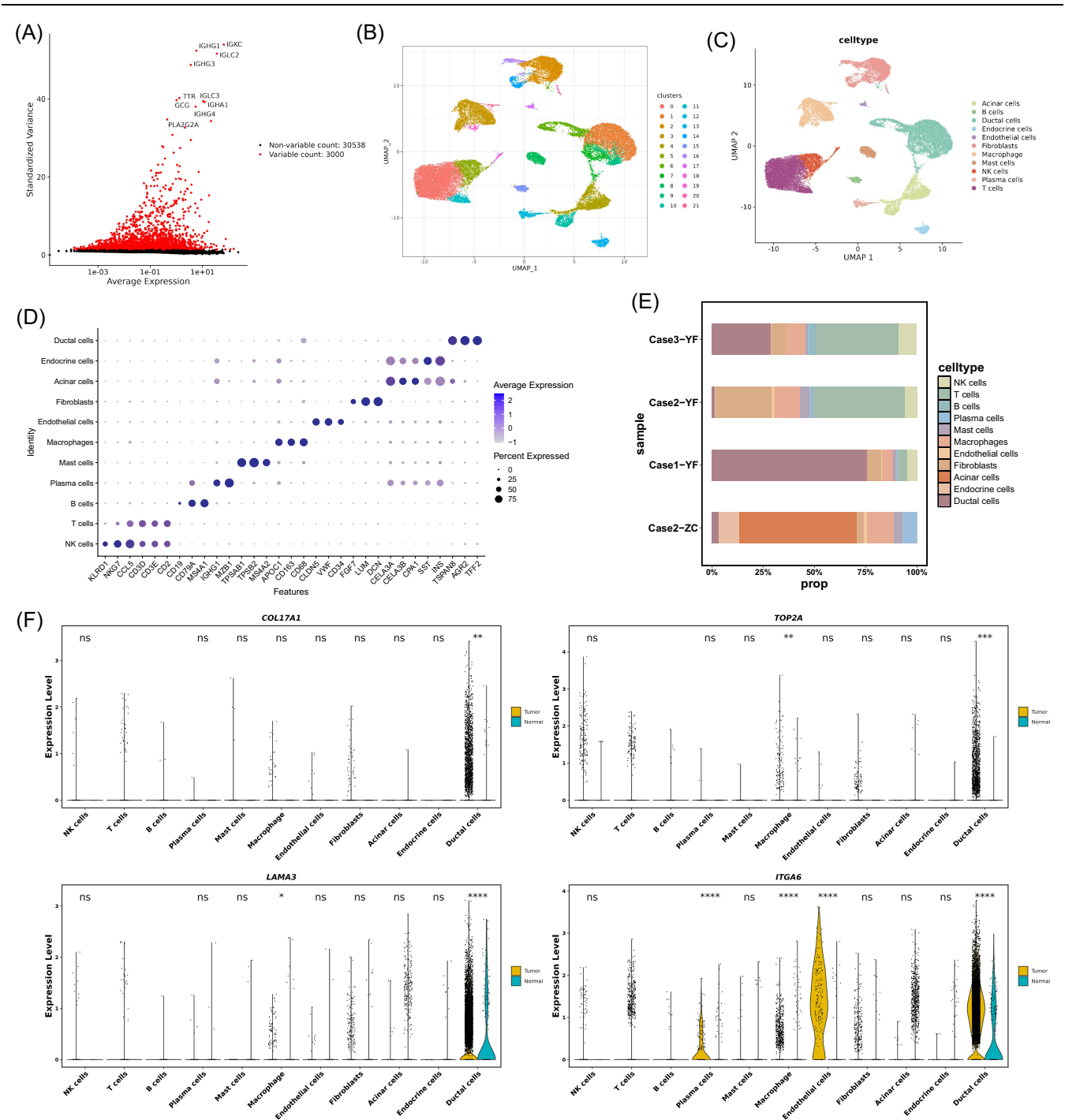


Figure 8. (A) Volcano plot of variable genes, with the x-axis representing gene expression levels and the y-axis showing standard deviations. Black dots denote non-highly variable genes, red dots indicate highly variable genes, and labeled gene names correspond to the top 10 highly variable genes. (B) UMAP plot of cell cluster classification, with the x-axis as UMAP 1 and the y-axis as UMAP 2, where colors represent different clusters. (C) Cell annotation plot, with the x-axis as UMAP 1 and the y-axis as UMAP 2, where colors indicate different annotated cell types. (D) Bubble plot of marker genes, with the x-axis as marker genes and the y-axis as annotated cell types. Colors represent average expression levels, with blue colors indicating higher expression, and dot sizes indicating the percentage of cells expressing the gene, with larger dots representing higher percentages. (E) Cell abundance plot, with the x-axis showing the percentage of all cells and the y-axis as samples, where colors represent different annotated cell types. (F) Plot of the expression levels of prognostic genes in various cell types within PDAC cancer tissue samples, with the x-axis as cell types and the y-axis as gene expression levels. Yellow represents the normal group, blue represents the tumor group, empty spaces indicate presence only in the tumor group, ns denotes no significant difference, * indicates $P < 0.05$, ** indicates $P < 0.01$, *** indicates $P < 0.001$, and **** indicates $P < 0.0001$.

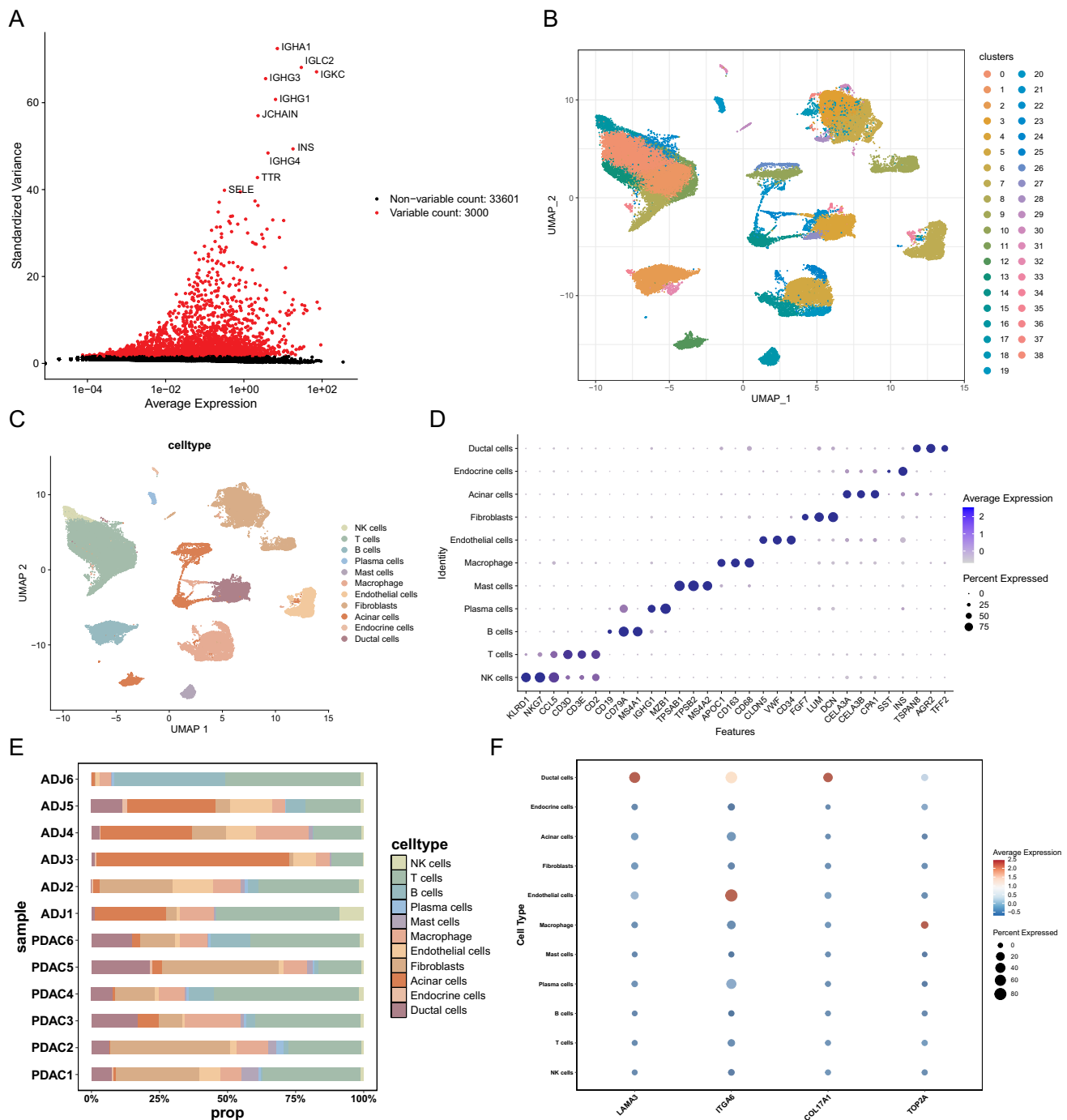


Figure 9. (A) Volcano plot of variable genes, with the x-axis representing gene expression levels and the y-axis showing standard deviations. Black dots denote non-highly variable genes, red dots indicate highly variable genes, and labeled gene names correspond to the top 10 highly variable genes. (B) UMAP plot of cell cluster classification, with the x-axis as UMAP 1 and the y-axis as UMAP 2, where colors represent different clusters. (C) Cell annotation plot, with the x-axis as UMAP 1 and the y-axis as UMAP 2, where colors indicate different annotated cell types. (D) Bubble plot of marker genes, with the x-axis as marker genes and the y-axis as annotated cell types. Colors represent average expression levels, with blue colors indicating higher expression, and dot sizes indicating the percentage of cells expressing the gene, with larger dots representing higher percentages. (E) Cell abundance plot, with the x-axis showing the percentage of all cells and the y-axis as samples, where colors represent different annotated cell types. (F) Dotplot of the expression levels of prognostic genes in various cell types within PDAC cancer tissue samples.

fibroblasts were evident across all tissue samples (Fig. 11B and C). These findings highlight the critical roles of ductal cells in PDAC pathogenesis.

Expression verification of prognostic SGRGs

Expression analysis of data from the PDAC-training set and GSE28735 revealed that LAMA3, ITGA6, COL17A1, and

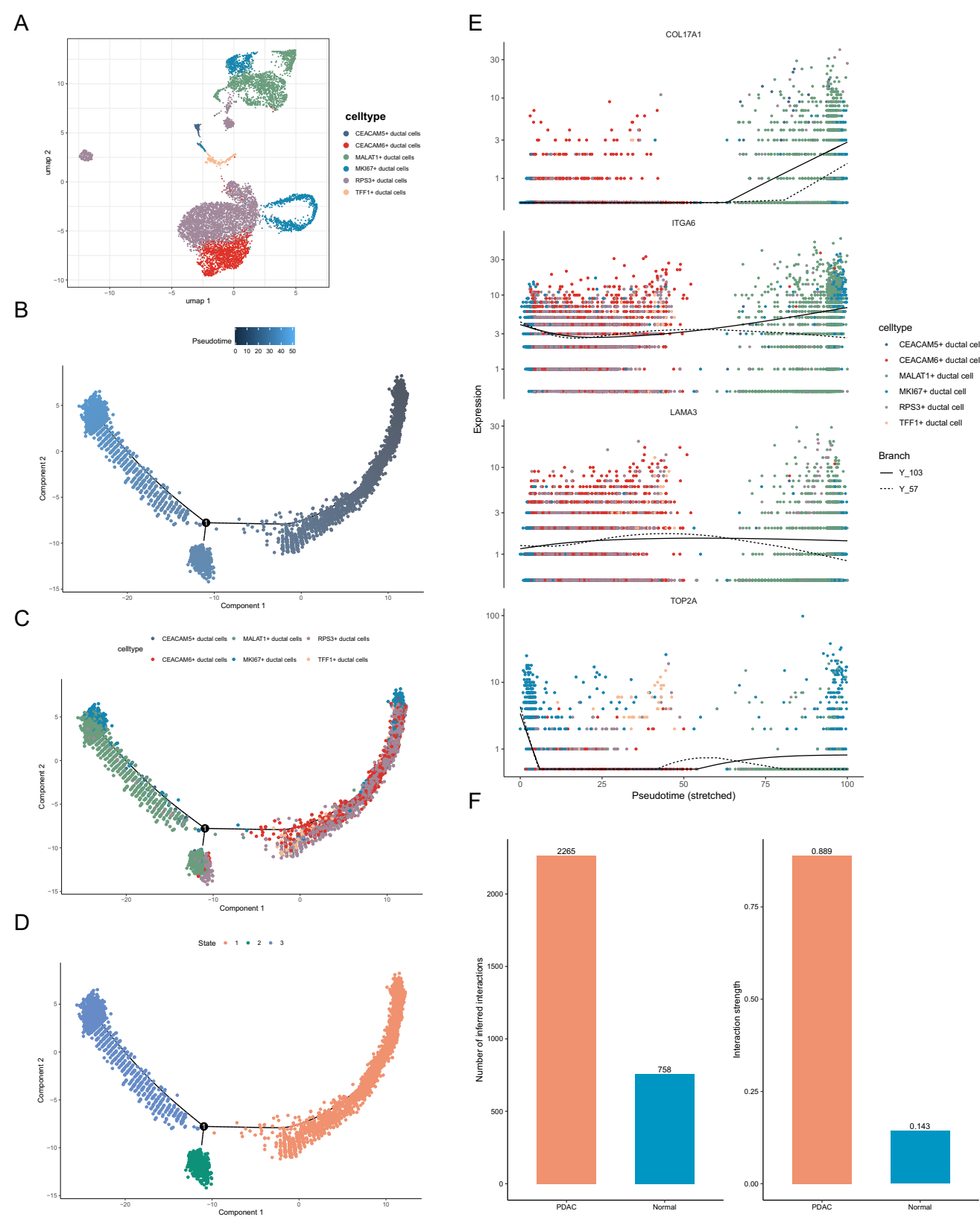


Figure 10. (A) UMAP plot of cell annotations for cell subtypes, with the x-axis as UMAP 1 and the y-axis as UMAP 2, where colors represent different annotated subpopulations of cells. (B and D) Pseudotime analysis plots for (B) pseudotime, (C) cell type, and (D) state, with the x-axis as component 1 and the y-axis as component 2. (E) Expression changes of prognostic genes LAMA3, ITGA6, COL17A1, and TOP2A during cell differentiation, with the x-axis as pseudotime values and the y-axis as gene expression levels, where colors indicate different annotated subpopulations of cells. The lines represent fitted lines for expression levels. (F) Bar plot showing the number and intensity of cell communication interactions.

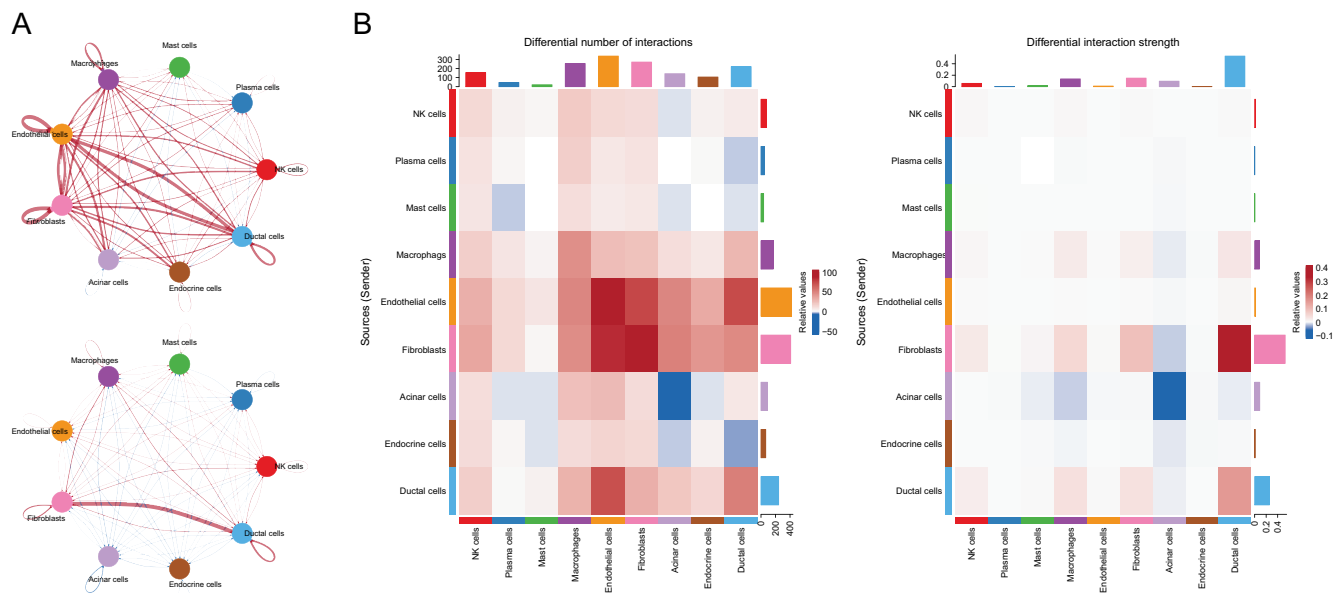


Figure 11. (A) Network plot comparing the number and intensity of interactions between different cell populations. (B and C) Heatmap of the number (B) and intensity (C) of cell communication interactions. The top-colored bar chart represents the sum of column values displayed in the heatmap (incoming signals). The right colored bar chart represents the sum of row values (outgoing signals). In the color bar, red (or blue) indicates an increase (or decrease) in signals in the second dataset compared to the first dataset.

TOP2A were significantly upregulated in PDAC tumor tissues ($P < 0.05$, Fig. 12A and B).

Discussion

This study utilized RNA sequencing and scRNA-seq data to unravel the complex landscape of PDAC, shedding light on cellular heterogeneity and its impact on tumor progression and patient outcomes. A PDAC risk model was successfully constructed using ML algorithms, accompanied by comprehensive analyses of immune cell infiltration patterns, drug sensitivity, and functional pathways, offering novel insights for the development of PDAC treatment strategies.

Biological function of prognostic SGRGs

A risk model was developed using data from a PDAC training set to explore the relationship between SGs and clinical outcomes in patients with PDAC. The model, based on prognostic SGRGs, demonstrated robust reliability and consistency, as validated by both the PDAC training set and the GSE28735 dataset. Four independent prognostic SGRGs – LAMA3, ITGA6, COL17A1, and TOP2A – were identified.

LAMA3, located at 18q11.2 on chromosome 18, encodes the laminin subunit alpha 3, a secreted protein within the laminin family^[44]. Laminin 332, of which LAMA3 is a key component, is critical to the basement membrane barrier and plays a significant role in tumor metastasis. The expression level of LAMA3 impacts the invasive and migratory capabilities of PDAC cells; higher LAMA3 expression facilitates basement membrane penetration, promoting tumor invasion and metastasis^[44,45]. The epithelial–mesenchymal transition (EMT), a process that endows epithelial cells with mesenchymal traits –

such as the loss of polarity and cell–cell junctions – enhances their migratory and invasive potential. In PDAC, LAMA3 promotes EMT by modulating key signaling pathways, thereby driving tumor progression^[46].

ITGA6, located at 2q31.1 on chromosome 2, encodes the integrin subunit alpha 6, which forms heterodimers with $\beta 1$ or $\beta 4$ subunits to interact with ECM proteins like laminin. In PDAC, this interaction may enhance tumor cell adhesion to the surrounding stroma, promoting cell migration and invasion and facilitating the breach of the basement membrane for tumor cell infiltration into adjacent tissues^[47]. The integrin $\alpha 6\beta 4$ complex is involved in angiogenesis-related signaling pathways, and in PDAC, ITGA6 may contribute to tumor angiogenesis by modulating these pathways, supplying nutrients, and offering metastatic routes for the tumor cells^[48,49]. Additionally, the interaction between ITGA6-mediated cells and the ECM activates intracellular signaling cascades that promote tumor cell proliferation, survival, and drug resistance. Through these pathways, ITGA6 likely supports tumor growth and progression in PDAC^[50].

COL17A1, located at 10q25.1 on chromosome 10, encodes the alpha 1 chain of type XVII collagen, a transmembrane protein that is a structural component of hemidesmosomes. It plays a critical role in mediating keratinocyte adhesion to the basement membrane at the dermal–epidermal junction^[51]. Elevated expression of COL17A1 has been linked to poor outcomes in pancreatic cancer, promoting tumor growth and metastasis^[52].

TOP2A, located at 17q21–q22 on chromosome 17, encodes DNA topoisomerase II α , a nuclear enzyme essential for DNA transcription and replication. TOP2A functions to control the local spatial conformation of DNA, participating in chromosome condensation, chromatid separation, and relieving torsional stress during DNA transcription and replication. Its role

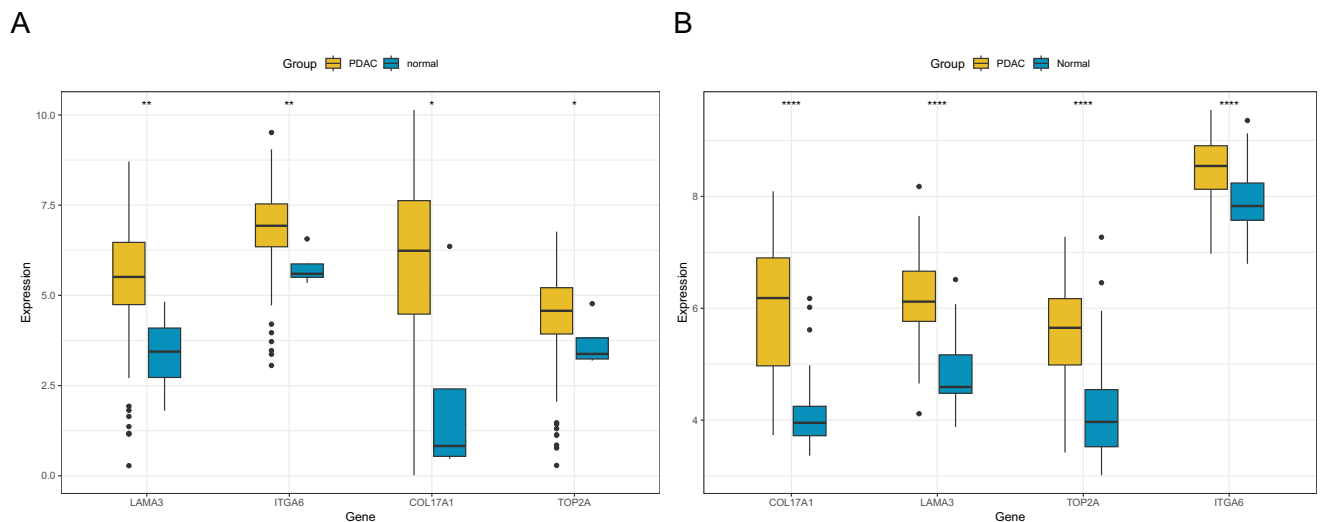


Figure 12. Representative plots of the expression levels of prognostic genes in cancerous and adjacent normal tissue samples from PDAC patients in the (A) training set and (B) testing set.

in DNA replication, transcription, and chromatin remodeling has implicated it in the progression of various cancers^[53]. Recent studies have noted a high rate of HER2 amplification in PDAC individuals with concurrent TOP2A amplification^[54]. TOP2A, along with its transcriptional activators SP1 and HMGB2, is overexpressed in human PDAC tissues, and knockdown of TOP2A sensitizes PDAC cells to chemotherapy^[55].

A risk model based on prognostic SGRGs was constructed and validated using one training set and three validation sets. Among the four datasets, the AUC value was below 0.6 at 1 year in one dataset, while exceeding 0.6 in the others, indicating relatively stable predictive performance of the model. Previous studies have similarly reported lower 1-year AUC values compared to those at 3 and 5 years^[56]. This discrepancy may result from the fact that, within the 1-year follow-up period, some “actual incident cases” could be misclassified as “non-incident” (false negatives) due to delayed screening (e.g., lack of annual check-ups) or diagnostic errors (e.g., false-negative biopsies). Additionally, certain “benign lesions” (e.g., breast fibroadenomas, hepatic hemangiomas) may be incorrectly identified as “suspected incident cases” (false positives). Moreover, our risk model represents an idealized framework based on bioinformatic analysis, and its clinical utility will require further validation and refinement through integration with clinical data in future studies.

Immune cell levels in high- and low-risk groups

Our findings revealed significant differences in the abundance of seven immune cell types between HRG and LRG in PDAC. Specifically, M0 macrophages and eosinophils demonstrated a significant positive correlation with all four prognostic SGRGs, while CD8 T cells showed a significant negative correlation with these SGRGs.

The PDAC stroma is heavily infiltrated by macrophages, which contribute to both immunosuppression within the tumor microenvironment and resistance to gemcitabine treatment^[57,58]. Moreover, macrophages can shield tumor cells from complement-dependent cytotoxicity^[58]. Early-stage M0

macrophages, characterized by high plasticity, can enhance anti-tumor responses by increasing TNF- α secretion^[59]. These macrophages may also interact with specific molecules on the surface of tumor cells, influencing tumor cell growth, proliferation, and invasiveness. For instance, M0 macrophages can bind to adhesion molecules, such as integrins, on the tumor cell surface, altering the ECM's adhesion properties and thereby impacting the migratory and invasive abilities of tumor cells^[60]. Furthermore, M0 macrophages may phagocytose tumor cell debris, release antigenic peptides, and activate specific antitumor immune responses.

Eosinophils, specialized white blood cells, play critical roles in immune defense, migrating to sites of infection and releasing toxic chemicals and proteins to eliminate harmful cells or pathogens^[61]. Elevated eosinophil levels can signal various health conditions, from allergic reactions and asthma to certain cancers^[62]. Retrospective research has demonstrated that a high peripheral blood eosinophil count in patients with PDAC correlates with prolonged survival^[63].

This study observed a significant reduction in CD8 T cell infiltration within the PDAC tumor microenvironment, likely due to the unique immunosuppressive conditions present. The highly fibrotic stroma in PDAC, where activated pancreatic stellate cells (PSCs) secrete ECM components and factors such as transforming growth factor- β ^[64], creates a physical barrier that impedes T cell infiltration. These factors may also induce functional exhaustion of CD8 T cells. Additionally, abnormal overexpression of immune checkpoint molecules, such as programmed death-ligand 1 (PD-L1)^[65], further suppresses the activation and proliferation of CD8 T cells. This not only exacerbates T cell exhaustion but also diminishes their tumor-killing capacity, contributing to immune evasion in PDAC.

These findings provide valuable insights into the immune cell profile differences between individuals at different risk levels, highlighting the importance of understanding the roles of specific immune cell populations in disease susceptibility and progression.

Gene mutations in high- and low-risk groups

By analyzing somatic mutation data from TCGA, this study investigated the mutational gene profiles of two PDAC risk groups. This analysis revealed that mutations in key genes, including KRAS, TP53, SMAD4, CDKN2A, and TTN, drive PDAC progression in these populations^[66–68]. Activating mutations in KRAS were found in 95% of pancreatic tumors^[69], while CDKN2A, TP53, and SMAD4 are frequently inactivated through methylation or homozygous deletion in 50–80% of cases^[70].

Drug sensitivity in high- and low-risk groups

Drug resistance remains a major therapeutic challenge in PDAC. Predicting drug sensitivity offers a way to personalize treatment by considering individual variations in chemotherapy response. Our findings suggest several potential therapeutic options for patients with PDAC, though the ultimate efficacy of these treatments will require validation in clinical trials.

Among the drugs with greater sensitivity in LRG, JAK1_8709 showed the smallest *p*-values. JAK1_8709 is a selective JAK1 kinase inhibitor that targets the Janus kinase–signal transducer and activator of transcription (JAK–STAT) pathway^[71]. In PDAC, this pathway is often hyperactivated by inflammatory cytokines such as IL-6, contributing to an immunosuppressive tumor microenvironment and promoting fibrosis^[72]. JAK1_8709 inhibits the JAK–STAT pathway, thereby impeding the formation of the tumor microenvironment, enhancing CD8 T cell infiltration, and potentially reversing PDAC's immune evasion phenotype.

For HRG, Acetalax showed the smallest *P*-values for drug sensitivity, although it was withdrawn in most countries in the early 1970s due to liver toxicity^[73]. The next most sensitive drug was SCH772984, a highly selective ERK1/2 kinase inhibitor that targets a critical node in the MAPK signaling pathway^[74]. KRAS mutations are common in PDAC, occurring in approximately 90% of cases, leading to hyperactivation of the MAPK pathway, which promotes tumor cell growth, proliferation, and survival^[75]. SCH772984 binds competitively to the ATP pocket of ERK, inhibiting its kinase activity and theoretically suppressing pancreatic cancer cell proliferation by blocking this pathway^[76,77]. The third most sensitive drug was trametinib, although studies have shown that combining trametinib with gemcitabine – both of which inhibit the MAPK signaling pathway (RAF>MEK1/2 > ERK1/2) – did not improve OS, progression-free survival, overall response rate, or duration of response in treatment-naïve patients with metastatic PDAC^[78].

Biological mechanisms underlying prognostic SGRGs in PDAC

GSEA results targeting prognostic SGRGs have revealed potential mechanisms underlying PDAC pathogenesis. The “Glucose Import” pathway may facilitate tumor cell proliferation by enhancing the Warburg effect, a metabolic phenomenon in which tumor cells primarily generate energy through glycolysis, even under normal oxygen conditions, leading to increased lactate production^[79]. In contrast to normal cells, which primarily rely on oxidative phosphorylation, cancer cells prefer glycolysis. The Warburg effect promotes tumor progression by providing cellular building blocks for rapid growth, creating an acidic

microenvironment, supporting immune evasion, and enhancing drug resistance^[80]. PDAC, characterized by abundant stromal components and a hypoxic, poorly vascularized tumor microenvironment^[81,82], increases tumor cell reliance on glycolytic pathways. The “Regulation of Immunoglobulin Production” pathway suggests that modulating B cell immune responses within the tumor microenvironment contributes to PDAC progression^[83]. Furthermore, “Renal System Vasculature Development” and “Blood Vessel Remodeling” emphasize the role of vascular remodeling in PDAC. IL-27 signaling within the “Interleukin 27 Mediated Signaling Pathway” may promote immune evasion by modulating the Th17/Treg balance^[84]. The “Protein Neddylation” modification pathway could influence the cell cycle by stabilizing oncogenic proteins^[85]. Additionally, metabolic pathways such as “NADPH Regeneration” and “Deoxyribonucleoside Triphosphate Biosynthetic Process” suggest that reprogramming nucleotide anabolic metabolism in PDAC may offer potential therapeutic targets^[86].

Among the SGRGs and their corresponding pathways, TOP2A is associated with the heparin biosynthetic process. In pancreatic cancer, aberrant expression of enzymes involved in heparin biosynthesis may impact processes such as tumor microenvironment modulation, cell proliferation, migration, and invasion. For example, heparanase, an enzyme involved in heparan sulfate (HS) degradation, is upregulated in pancreatic cancer and promotes tumor cell migration and invasion. Additionally, changes in the expression of enzymes like HS sulfotransferases can alter HS structure and function, influencing PDAC progression. HS3ST3B1, for example, can promote EMT in pancreatic cancer cells by activating the transcription factor SNAIL^[87]. The upregulation of TOP2A affects the heparin biosynthetic process, thereby influencing tumor cell migration and PDAC progression. Meanwhile, LAMA3 is enriched in the sulfation pathway, and studies have shown that inhibiting the tyrosine sulfation axis of SLC35B2–TPST2 can reduce growth and metastasis in PDAC^[88].

The role of ductal cells in the progression of PDAC

Previous single-cell analyses have highlighted the critical roles of ductal cells in PDAC. This cancer is believed to arise from the malignant transformation of pancreatic ductal epithelial cells, closely associated with both genetic and epigenetic abnormalities in these cells. During the precancerous stage, ductal cells often harbor oncogenic KRAS mutations, present in approximately 90% of PDAC cases. These mutations lead to the persistent activation of the MAPK/ERK signaling pathway^[89], driving cell proliferation, inhibiting apoptosis, and inducing metabolic reprogramming. This provides energy and biosynthetic precursors necessary for the rapid growth of tumor cells. The PDAC tumor microenvironment is also marked by a dense stromal component, with ductal cells recruiting fibroblasts, immune cells, and vascular endothelial cells through the secretion of cytokines and chemokines, thus promoting a protumorigenic environment^[7]. Additionally, ductal cells secrete the sonic hedgehog (SHH) protein, which activates PSCs and induces collagen deposition, resulting in stromal stiffening that impedes drug penetration^[90]. Notably, the expression dynamics observed in the quasitemporal analysis do not indicate causal regulation. Given the strong interaction between ductal cells and fibroblasts, it can be hypothesized that the prominent expression of

certain biomarkers might facilitate this intercellular communication. For instance, TOP2A, in synergy with TGF- β secreted by cancer-associated fibroblasts (CAFs), promotes remodeling of the ductal cell ECM, creating a physical barrier that hinders immune cell infiltration and increases chemotherapy resistance^[91,92]. Thus, the upregulated expression of TOP2A in this study likely contributes to ECM remodeling in ductal cells, further obstructing immune cell infiltration and enhancing resistance to chemotherapy.

Limitations and future perspectives

In summary, this study utilized SGs as a foundation to thoroughly investigate the molecular mechanisms of prognostic SGRGs in PDAC. The resulting risk model demonstrated robust performance across multiple datasets, exhibiting strong prognostic predictive capabilities and potential clinical applications. Moreover, the analysis of drug sensitivity provided new targets for PDAC immunotherapy and drug development, while single-cell analysis results offered a more comprehensive understanding of PDAC pathogenesis.

Despite these strengths, the study has certain limitations. The current analysis of SG heterogeneity is based on bioinformatics predictions and lacks direct experimental validation, such as examining the assembly patterns of SGs (e.g., *via* immunofluorescence staining) across different PDAC subpopulations. Future research will focus on using PDAC tissue microarrays and immunofluorescence colocalization techniques to validate the differences in SG assembly among various cell populations and samples with distinct pathological characteristics. Additionally, *in vitro* cell models (poorly differentiated and well-differentiated PDAC cell lines) will be employed to explore the regulatory mechanisms of SG heterogeneity. Furthermore, future studies should explore the interactions between SGRGs and other metabolic and signaling pathways in PDAC. Leveraging single-cell data to investigate the differentiation processes of key cell types, such as ductal cells, and elucidating their roles in PDAC, along with the molecular mechanisms of cellular differentiation, offers a promising direction for further investigation. Finally, as this study is based on computational predictions, it lacks validation through functional experiments and clinical sample testing. Future research will continue to prioritize experimental validation to facilitate the clinical translation of these findings.

Conclusion

This study utilized ML approaches to develop a PDAC risk model that incorporates four prognostic SGRGs (LAMA3, ITGA6, COL17A1, and TOP2A). This model independently predicts the prognosis of patients with PDAC, demonstrating superior accuracy in OS prediction compared to traditional factors such as age, gender, and N stage. The calibration curve confirmed a high degree of agreement between the predicted and actual survival rates. This scoring system offers a personalized survival prediction tool for clinicians and patients, supporting the development of more effective treatment strategies. Additionally, this study highlighted significant differences in the tumor immune microenvironment between high- and low-risk groups, including variations in Naive B cells, Memory B cells, CD8 T cells, Monocytes, M0 Macrophages, M1 Macrophages, and Eosinophils. The ESTIMATE algorithm revealed significant differences in immune, stromal, and

combined scores between these risk groups. Furthermore, analysis of relevant single-cell datasets identified 11 distinct cell types, including ductal cells, mast cells, fibroblasts, macrophages, NK cells, T cells, plasma cells, acinar cells, endothelial cells, B cells, and endocrine cells. Ductal cells emerged as a key cell cluster. This study further explored cellular state transitions and cell-cell communication within ductal cells, providing a deeper understanding of PDAC pathogenesis.

Ethical approval

Not applicable. The data used in this study were all sourced from freely accessible public databases.

Consent

Not applicable. This study did not involve patients or volunteers.

Sources of funding

National Natural Science Foundation of China under Grant No. 42207324.

Author contributions

L.W. and M.G. designed the study. L.W. and L.Z. conducted data analysis. L.Z. visualized the data. L.W. and L.Z. drafted the manuscript. M.G., Y.W., and Y.J. reviewed and revised the manuscript.

Conflicts of interest disclosure

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Guarantor

Miao Guan.

Research registration unique identifying number (UIN)

Not applicable. This study does not involve human subjects.

Provenance and peer review

This paper is not commissioned, externally peer-reviewed.

Data availability statement

Data will be available upon reasonable request. Publicly available datasets were analyzed in this study. These data can be found at <https://www.cancer.gov/ccg/research/genome-sequencing/tcga>; <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE28735>; and <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE197177>.

Acknowledgements

All individuals who contributed to this study are included in the list of authors. We sincerely acknowledge the GEO and TCGA databases, as well as the uploaders of the datasets.

References

- [1] Wu Y, Zhang C, Huang J, *et al.* Integrated analysis of scRNA-seq and bulk RNA-seq data identifies BHLHE40 as a key gene in pancreatic cancer progression and gemcitabine resistance. *Semin Oncol* 2025;52:152338.
- [2] Wang W, Hu K, Xue J, *et al.* In vivo FAP-CAR macrophages enhance chemotherapy and immunotherapy against pancreatic cancer by removing the fibrosis barrier. *J Control Release* 2025;384:113888.
- [3] Murakawa M, Kawahara S, Takahashi D, *et al.* Risk factors for early recurrence in patients with pancreatic ductal adenocarcinoma who underwent curative resection. *World J Surg Oncol* 2023;21:263.
- [4] Hu ZI, O'Reilly EM: therapeutic developments in pancreatic cancer. *Nat Rev Gastroenterol Hepatol* 2024;21:7–24.
- [5] Murray K, Oldfield L, Stefanova I, *et al.* Biomarkers, omics and artificial intelligence for early detection of pancreatic cancer. *Semin Cancer Biol* 2025;111:76–88.
- [6] Bear AS, Vonderheide RH, O'Hara MH: challenges and Opportunities for Pancreatic Cancer Immunotherapy. *Cancer Cell* 2020;38:788–802.
- [7] Sherman MH, Beatty GL. Tumor microenvironment in pancreatic cancer pathogenesis and therapeutic resistance. *Annu Rev Pathol* 2023;18:123–48.
- [8] Aramburu-Núñez M, Custodia A, Pérez-Mato M, *et al.* Stress granules and acute ischemic stroke: beyond mRNA translation. *Int J Mol Sci* 2022;23:3747.
- [9] Fonteneau G, Redding A, Hoag-Lee H, *et al.* Stress granules determine the development of obesity-associated pancreatic cancer. *Cancer Discov* 2022;12:1984–2005.
- [10] Schneider C, Schneider G. Stress granules-membraneless organelles as therapeutic targets in pancreatic cancer. *EMBO Mol Med* 2024;16:429–31.
- [11] Li Z, Ge Y, Dong J, *et al.* BZW1 facilitates glycolysis and promotes tumor growth in pancreatic ductal adenocarcinoma through potentiating eIF2 α phosphorylation. *Gastroenterology* 2022;162:1256–1271.e1214.
- [12] Guo S, Liu X, Zhang J, *et al.* Integrated analysis of single-cell RNA-seq and bulk RNA-seq unravels T cell-related prognostic risk model and tumor immune microenvironment modulation in triple-negative breast cancer. *Comput Biol Med* 2023;161:107066.
- [13] Ma P, Amemiya HM, He LL, *et al.* Bacterial droplet-based single-cell RNA-seq reveals antibiotic-associated heterogeneous cellular states. *Cell* 2023;186:877–891.e814.
- [14] Ge H, Wolters-Eisfeld G, Hackert T, Li Y, Güngör C. Development of a hypoxia-responsive macrophage prognostic model using single-cell and bulk RNA sequencing in pancreatic cancer. *PLoS One* 2025;20:e0322618.
- [15] Zaccaria GM, Altini N, Mongelli V, Marino F, Bevilacqua V. Development and validation of a machine learning prognostic model based on an epigenomic signature in patients with pancreatic ductal adenocarcinoma. *Int J Med Inform* 2025;199:105883.
- [16] Dusetti N, Bachet JB, Chanez B, *et al.* Medical management of pancreatic cancer: from personalization to broadening treatment strategies. *Cancer Treat Rev* 2025;138:102973.
- [17] Liu W, Li J, Yuan X, *et al.* A novel deep learning-based pathomics score for prognostic stratification in pancreatic ductal adenocarcinoma. *Pancreas* 2025;54:e430–e441.
- [18] Sauerbrei W, Taube SE, McShane LM, Cavenagh MM, Altman DG. Reporting recommendations for tumor marker prognostic studies (REMARK): an abridged explanation and elaboration. *J Natl Cancer Inst* 2018;110:803–11.
- [19] Agha RA, Mathew G, Rashid R, *et al.* Group T: transparency in the Reporting of Artificial Intelligence – the TITAN Guideline. *Prem J Sci* 2025;10:100082.
- [20] Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 2014;15:550.
- [21] Gu Z, Eils R, Schlesner M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* 2016;32:2847–49.
- [22] Gustavsson EK, Zhang D, Reynolds RH, Garcia-Ruiz S, Ryten M. ggtranscript: an R package for the visualization and interpretation of transcript isoforms using ggplot2. *Bioinformatics* 2022;38:3844–46.
- [23] Zheng Y, Gao W, Zhang Q, *et al.* Ferroptosis and autophagy-related genes in the pathogenesis of ischemic cardiomyopathy. *Front Cardiovasc Med* 2022;9:906753.
- [24] Yu G, Wang LG, Han Y, He QY. clusterProfiler: an R package for comparing biological themes among gene clusters. *Omics* 2012;16:284–87.
- [25] Lei J, Qu T, Cha L, *et al.* Clinicopathological characteristics of pheochromocytoma/paraganglioma and screening of prognostic markers. *J Surg Oncol* 2023;128:510–18.
- [26] Liu H, Zhang W, Zhang Y, *et al.* Mime: a flexible machine-learning framework to construct and visualize models for clinical characteristics prediction and feature selection. *Comput Struct Biotechnol J* 2024;23:2798–810.
- [27] Zhang N, Zhang H, Wu W, *et al.* Machine learning-based identification of tumor-infiltrating immune cell-associated lncRNAs for improving outcomes and immunotherapy responses in patients with low-grade glioma. *Theranostics* 2022;12:5931–48.
- [28] Luo C, Zhang J, Bo L, *et al.* Construction of a ceRNA-based lncRNA-mRNA network to identify functional lncRNAs in premature ovarian insufficiency. *Front Genet* 2022;13:956805.
- [29] Heagerty PJ, Lumley T, Pepe MS. Time-dependent ROC curves for censored survival data and a diagnostic marker. *Biometrics* 2000;56:337–44.
- [30] Liu TT, Li R, Huo C, *et al.* Identification of CDK2-related immune forecast model and ceRNA in lung adenocarcinoma, a pan-cancer analysis. *Front Cell Dev Biol* 2021;9:682002.
- [31] Sachs MC. plotROC: a Tool for Plotting ROC Curves. *J Stat Softw* 2017;79:1–9.
- [32] Newman AM, Liu CL, Green MR, *et al.* Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods* 2015;12:453–57.
- [33] Orifjon S, Jammатов J, Sousa C, Barros R, Vasconcelos O, Rodrigues P. Translation and adaptation of the adult developmental coordination disorder/dyspraxia checklist (ADC) into Asian Uzbekistan. *Sports (Basel)* 2023;11:135.
- [34] Yoshihara K, Shahmoradgoli M, Martínez E, *et al.* Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* 2013;4:2612.
- [35] Zhuo Z, Lin H, Liang J, *et al.* Mitophagy-related gene signature for prediction prognosis, immune scenery, mutation, and chemotherapy response in pancreatic cancer. *Front Cell Dev Biol* 2021;9:802528.
- [36] Maeser D, Gruener RF, Huang RS. OncoPredict: an R package for predicting in vivo or cancer patient drug response and biomarkers from cell line screening data. *Brief Bioinform* 2021;22:bbab260.
- [37] Hyeon DY, Nam D, Han Y, *et al.* Proteogenomic landscape of human pancreatic ductal adenocarcinoma in an Asian population reveals tumor cell-enriched and immune-rich subtypes. *Nat Cancer* 2023;4:290–307.
- [38] Mayakonda A, Lin DC, Assenov Y, Plass C, Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018;28:1747–56.
- [39] Hao Y, Hao S, Andersen-Nissen E, *et al.* Integrated analysis of multi-modal single-cell data. *Cell* 2021;184:3573–3587.e3529.
- [40] Zhang S, Fang W, Zhou S, *et al.* Single cell transcriptomic analyses implicate an immunosuppressive tumor microenvironment in pancreatic cancer liver metastasis. *Nat Commun* 2023;14:5123.
- [41] Wang X, Wang G, Xu Q, *et al.* Potential tumor-specific antigens and immune landscapes identification for mRNA vaccine in thyroid cancer. *Front Oncol* 2024;14:1480028.
- [42] Jin S, Guerrero-Juarez CF, Zhang L, *et al.* Inference and analysis of cell-cell communication using CellChat. *Nat Commun* 2021;12:1088.
- [43] Chatsirisupachai K, Lager C, de Magalhães JP. Age-associated differences in the cancer molecular landscape. *Trends Cancer* 2022;8:962–71.
- [44] Xing Y, Jing X, Qing G, Jiang Y. Correlation of laminin subunit alpha 3 expression in pancreatic ductal adenocarcinoma with tumor liver metastasis and survival. *Radiol Oncol* 2024;58:234–42.
- [45] Yang C, Liu Z, Zeng X, *et al.* Evaluation of the diagnostic ability of laminin gene family for pancreatic ductal adenocarcinoma. *Aging (Albany NY)* 2019;11:3679–703.

- [46] Huang C, Chen J. Laminin-332 mediates proliferation, apoptosis, invasion, migration and epithelial-to-mesenchymal transition in pancreatic ductal adenocarcinoma. *Mol Med Rep* 2021;23:11.
- [47] Asada T, Nakahata S, Fauzi YR, *et al.* Integrin $\alpha 6A$ (ITGA6A)-type splice variant in extracellular vesicles has a potential as a novel marker of the early recurrence of pancreatic cancer. *Anticancer Res* 2022;42:1763–75.
- [48] Robinson CJ, Thiagarajan L, Maynard R, *et al.* Release of miR-29 target laminin C2 improves skin repair. *Am J Pathol* 2024;194:195–208.
- [49] Samarelli AV, Riccitelli E, Bizzozero L, *et al.* Neuroligin 1 induces blood vessel maturation by cooperating with the $\alpha 6$ integrin. *J Biol Chem* 2014;289:19466–76.
- [50] Li J, Tan W, Peng L, *et al.* Integrative analysis of gene expression profiles reveals specific signaling pathways associated with pancreatic duct adenocarcinoma. *Cancer Commun (Lond)* 2018;38:13.
- [51] Shi R, Zhang N, Li H, *et al.* Cancer-associated fibroblast-derived COL17A1 promotes gemcitabine resistance and tumorigenesis in pancreatic cancer cells by interacting with ACTN4. *Discov Oncol* 2025;16:118.
- [52] Kozawa K, Sekai M, Ohba K, *et al.* The CD44/COL17A1 pathway promotes the formation of multilayered, transformed epithelia. *Curr Biol* 2021;31:3086–3097.e3087.
- [53] Chen T, Sun Y, Ji P, Kopetz S, Zhang W. Topoisomerase II α in chromosome instability and personalized cancer therapy. *Oncogene* 2015;34:4019–31.
- [54] Heestand GM, Schwaederle M, Gatalica Z, Arguello D, Kurzrock R. Topoisomerase expression and amplification in solid tumours: analysis of 24,262 patients. *Eur J Cancer* 2017;83:80–87.
- [55] Tanaka T, Okada R, Hozaka Y, *et al.* Molecular Pathogenesis of Pancreatic Ductal Adenocarcinoma: impact of miR-30c-5p and miR-30c-2-3p Regulation on Oncogenic Genes. *Cancers (Basel)* 2020;12:2731.
- [56] Peng T, Sun F, Yang JC, *et al.* Novel lactylation-related signature to predict prognosis for pancreatic adenocarcinoma. *World J Gastroenterol* 2024;30:2575–602.
- [57] Spek CA, Abernethy HL, Duitman J. Macrophage C/EBP δ Drives Gemcitabine, but Not 5-FU or paclitaxel, resistance of pancreatic cancer cells in a deoxycytidine-dependent manner. *Biomedicines* 2022;10:219.
- [58] Yang J, Li Y, Sun Z, Zhan H. Macrophages in pancreatic cancer: an immunometabolic perspective. *Cancer Lett* 2021;498:188–200.
- [59] Tekin C, Abernethy HL, Bijlsma MF, Spek CA. Early macrophage infiltrates impair pancreatic cancer cell growth by TNF- α secretion. *BMC Cancer* 2020;20:1183.
- [60] Ju Y, Xu D, Liao MM, *et al.* Barriers and opportunities in pancreatic cancer immunotherapy. *Npj Precision Oncology* 2024;8:199.
- [61] Wen T, Rothenberg ME. The Regulatory function of EOSINOPHILS. *Microbiol Spectr* 2016;4:10–128.
- [62] Ramirez GA, Yacoub MR, Ripa M, *et al.* Eosinophils from physiology to disease: a comprehensive review. *Biomed Res Int* 2018;2018:9095275.
- [63] Ciesielski W, Kupryś-Lipińska I, Kumor-Kisielewska A, *et al.* Peripheral eosinophil count may be the prognostic factor for overall survival in patients with pancreatic ductal adenocarcinoma undergoing surgical treatment. *Biomedicines* 2024;12:2596.
- [64] Zhao Y, Li Y, Zou J, *et al.* Low-dose arsenic trioxide inhibits pancreatic stellate cell activation via LOXL3 expression to enhance immunotherapy in pancreatic cancer. *Br J Cancer* 2024;131:1928–41.
- [65] Li X, Gulati M, Larson AC, *et al.* Immune checkpoint blockade in pancreatic cancer: trudging through the immune desert. *Semin Cancer Biol* 2022;86:14–27.
- [66] Li Z, Ren H, Zhang S, *et al.* PD-L1 levels, TP53 mutation profiles, and survival outcomes in pancreatic cancer differ by immune-nutritional status. *World J Surg Oncol* 2025;23:174.
- [67] Toriyama K, Masago K, Shibata N, *et al.* Clinicopathological and molecular characterization of KRAS wild-type pancreatic ductal adenocarcinomas reveals precursor lesions with oncogenic mutations and fusions in RAS pathway genes. *J Pathol* 2025;266:337–51.
- [68] Yousef M, Yousef A, Hurd MW, *et al.* KRAS mutation detection by liquid biopsy for pancreatic ductal adenocarcinoma. *J Hematol Oncol* 2025;18:44.
- [69] Network CGAR. Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell* 2017;32:185–203.e113.
- [70] Kleeff J, Korc M, Apte M, *et al.* Pancreatic cancer. *Nat Rev Dis Primers* 2016;2:16022.
- [71] Su H, Chen Y, Wang W. Novel prognostic model of complement and coagulation cascade-related genes correlates with immune environment and drug sensitivity in hepatocellular carcinoma. *Heliyon* 2024;10:e38230.
- [72] Shrestha H, Rädler PD, Dennaoui R, *et al.* The Janus kinase 1 is critical for pancreatic cancer initiation and progression. *Cell Rep* 2024;43:114202.
- [73] Reinhold WC, Marangoni E, Elloumi F, *et al.* Acetalax and bisacodyl for the treatment of triple-negative breast cancer: a combined molecular and preclinical study. *Cancer Res Commun* 2025;5:375–88.
- [74] Itoh M, Tohda S. Effects of ERK1/2 Inhibitors on the Growth of Acute Leukemia Cells. *Anticancer Res* 2024;44:5263–70.
- [75] Sudhakar N, Yan L, Qiryaqos F, *et al.* The SOS1 Inhibitor MRTX0902 Blocks KRAS activation and demonstrates antitumor activity in cancers dependent on KRAS nucleotide loading. *Mol Cancer Ther* 2024;23:1418–30.
- [76] Bye BA, Jack JL, Pierce A, *et al.* Combined Omipalisib and MAPK inhibition suppress PDAC Growth. *Cancers (Basel)* 2025;17:1152.
- [77] Zhao B, Fang R, Schürmann H, *et al.* PLK1 blockade enhances the anti-tumor effect of MAPK inhibition in pancreatic ductal adenocarcinoma. *Cell Rep* 2025;44:115541.
- [78] Silvis MR, Silva D, Rohweder R, *et al.* MYC-mediated resistance to trametinib and HCQ in PDAC is overcome by CDK4/6 and lysosomal inhibition. *J Exp Med* 2023;220:e20221524.
- [79] Li Q, Liu Y, Wang Z, Li Y, He L, Li J. Knockdown of MACC1 expression attenuates colorectal cancer cell glucose metabolism by suppressing GLUT4 membrane translocation. *Cancer Biomark* 2025;42:1875859 2251332426.
- [80] Barba I, Carrillo-Bosch L, Seoane J. Targeting the Warburg Effect in Cancer: where Do We Stand? *Int J Mol Sci* 2024;25:3142.
- [81] Chen PC, Ning Y, Li H, *et al.* Targeting ONECUT3 blocks glycolytic metabolism and potentiates anti-PD-1 therapy in pancreatic cancer. *Cell Oncol Dordr* 2024;47:81–96.
- [82] Zhu L, Li B, Li R, *et al.* METTL3 suppresses pancreatic ductal adenocarcinoma progression through activating endogenous dsRNA-induced anti-tumor immunity. *Cell Oncol Dordr* 2023;46:1529–41.
- [83] Ni P, Li L, Du K, *et al.* Unveiling the immunological terrain of pancreatic ductal adenocarcinoma: strategies to prompt immunotherapy from Mendelian randomization. *Discov Oncol* 2025;16:613.
- [84] Yao G, Qi J, Liang J, *et al.* Mesenchymal stem cell transplantation alleviates experimental Sjögren's syndrome through IFN- β /IL-27 signaling axis. *Theranostics* 2019;9:8253–65.
- [85] Qin X, Dang W, Yang X, Wang K, Kebreab E, Lyu L. Neddylation inactivation affects cell cycle and apoptosis in sheep follicular granulosa cells. *J Cell Physiol* 2022;237:3278–91.
- [86] Song Q, Zhang K, Sun T, Xu C, Zhao W, Zhang Z. Knockout of ENO1 leads to metabolism reprogramming and tumor retardation in pancreatic cancer. *Front Oncol* 2023;13:1119886.
- [87] Ma SN, Mao ZX, Wu Y, *et al.* The anti-cancer properties of heparin and its derivatives: a review and prospect. *Cell Adh Migr* 2020;14:118–28.
- [88] Cai X, Li S, Zeng X, *et al.* Inhibition of the SLC35B2-TPST2 axis of tyrosine sulfation attenuates the growth and metastasis of pancreatic ductal adenocarcinoma. *Cell Mol Gastroenterol Hepatol* 2023;16:473–95.
- [89] Flores LF, Marks DL, Vera RE, *et al.* Emerin is an effector of oncogenic KRAS-driven nuclear dynamics in pancreatic cancer. *JCI Insight* 2025;10:e187799.
- [90] Tsang SW, Zhang H, Lin C, *et al.* Rhein, a natural anthraquinone derivative, attenuates the activation of pancreatic stellate cells and ameliorates pancreatic fibrosis in mice with experimental chronic pancreatitis. *PLoS One* 2013;8:e82201.
- [91] Wang X, Wang J, Lyu L, Gao X, Cai Y, Tang B. Oncogenic role and potential regulatory mechanism of topoisomerase II α in a pan-cancer analysis. *Sci Rep* 2022;12:11161.
- [92] Wilczyński B, Dąbrowska A, Kulbacka J, Baczyńska D. Chemoresistance and the tumor microenvironment: the critical role of cell-cell communication. *Cell Commun Signal* 2024;22:486.