

Brief Communication

Platelet-related gene risk score: a predictor for pancreatic cancer microenvironmental signature, chemosensitivity and prognosis

Meng-Si Hu¹, Ming Jiang², Ying-Jian Wang¹, Shou-Fang Xu¹, Fei-Yu Jiang¹, Ye-Tao Han¹, Zhi-Wei Liu¹, Hong Yu²

¹Department of Blood Transfusion, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, 3th East Qingchun Road, Hangzhou 310016, Zhejiang, P. R. China; ²Department of General Surgery, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310016, Zhejiang, P. R. China

Received September 3, 2023; Accepted November 21, 2023; Epub December 15, 2023; Published December 30, 2023

Abstract: Recent studies have indicated that platelets may play a role in the advancement of pancreatic cancer by supporting tumor growth and increasing resistance to chemotherapy. This study aims to develop a prognostic model for pancreatic cancer using a platelet-related gene risk score. Prognostic platelet-related genes (PRGs) were identified from public databases and analyzed using cluster analysis. We investigated the microenvironment signatures and gene mutation patterns across different PRG-based molecular subtypes of pancreatic cancer. A prognostic model based on PRGs was developed using LASSO-Cox Regression Analysis. Additionally, we examined the correlation between the risk score and tumor clinical characteristics, as well as drug sensitivity. Two molecular subtypes, cluster C1 and C2, were identified. Cluster C2 was associated with a poorer prognosis compared to Cluster C1. The C1 group exhibited higher scores for activated CD8+ T cells, central memory CD4+ T cells, and natural killer T cells. The C2 group demonstrated a higher frequency of gene mutations. We established and validated a novel prognostic prediction model and platelet-related gene risk score for pancreatic cancer. The risk score was positively correlated with T stage, N stage, and tumor grade, and it presented a significant prognostic value compared to other clinical factors. In conclusion, a novel prognostic prediction model focusing on platelet involvement in pancreatic cancer has been developed, offering potential benefits for future drug therapies and clinical prognostic assessments.

Keywords: Platelets, pancreatic cancer, tumor microenvironment signature, chemosensitivity, gene signature, prognosis

Introduction

Pancreatic adenocarcinoma (PAAD) ranks as the fourth leading cause of cancer-related deaths in the USA and is responsible for approximately 227,000 deaths annually worldwide [1]. Pancreatic ductal adenocarcinoma (PDAC), constituting 80-90% of pancreatic cancer cases, is characterized by rapid progression, a poor prognosis, and an overall 5-year survival rate of less than 10% [2]. Currently, surgery is the only method that significantly improves postoperative survival and extends life expectancy in pancreatic cancer patients. However, surgery is an option for only about 20% of patients, and even then, the rate of postoperative recurrence remains high. Despite these challenges, effective biomarkers to predict

postoperative survival in PDAC patients are still lacking. Consequently, identifying predictive indicators remains a critical and urgent clinical challenge.

In the past decade, mounting evidence has indicated that platelets play a pivotal role in the proliferation, metastasis, and angiogenesis of cancer, potentially accelerating pancreatic cancer progression by fostering tumor growth and increasing chemotherapy resistance [3, 4]. Platelets are often referred to as the “little helpers” of tumors. Recent research has examined the transcriptomes of human PDAC circulating tumor cells (CTCs), primary tumors, and metastatic lesions at a single-cell level. This study found significant upregulation of platelet-related genes in CTCs, suggesting a potential link

Platelet genes predict pancreatic cancer prognosis

between tumors and platelets. There is a notable association between CTCs and platelet aggregation and activation, with CTC metastasis likely being influenced by platelets [5]. Further research has revealed that patients with advanced pancreatic cancer exhibiting higher platelet levels in their blood tend to have poorer prognoses and diminished responses to chemotherapy. Platelets may release microparticles that contribute to tumor growth and metastasis in pancreatic cancer [6]. Thus, the identification of novel biomarkers for predicting the prognosis of PAAD patients is essential.

In our current study, we strive to establish a predictive model for the clinical prognosis of pancreatic cancer. This model is developed by analyzing differentially expressed genes associated with platelets, tumor microenvironment features, and molecular mutations. It incorporates a risk assessment based on platelet-related prognostic genes, which correlates with clinical characteristics and drug sensitivity.

Materials and methods

Study design and data collection

The flow chart of our present study was demonstrated in [Figure S1](#). We sourced the clinical phenotype data for pancreatic cancer from the UCSC Xena database (<http://xena.ucsc.edu/>), excluding samples lacking survival time and status. We included patients with a documented survival time of at least 0 day. The expression profile data for TCGA-PAAD was retrieved from The Cancer Genome Atlas (TCGA) database (<https://www.cancer.gov/ccg/>) for 176 tumor samples. As a control group, we obtained corresponding normal pancreatic tissue sequencing data from the Genotype-Tissue Expression (GTEx) database, which consisted of 165 samples. Additionally, we collected four GEO datasets that included complete survival time and status: GSE57495 [7] (63 cases), GSE62452 [8] (66 cases), GSE71729 [9] (123 cases), and GSE85916 (79 cases). Both TCGA and GTEx expression data were measured in fragments per kilobase of transcript per million mapped reads (FPKM) format. TCGA biolinks [10] (version 2.24.3) was used to download TCGA-PAAD gene SNV mutation information calculated by the Mutect2 software. The homologous recombination defects (HRD), fraction altered (FA), nonsilent mutation rate (NSM),

number of altered segments (NAS), Aneuploidy Score, and silent mutation rate (SSM) data for each TCGA sample were sourced from previous studies [11].

Identification of platelet-related differential genes

We collected research data on 300 platelet-related genes (PRGs) associated with platelet biological function [12], and screened for differentially expressed platelet-related genes (DEPRGs) in the TCGA-PAAD cohort using the limma package (version 3.50.3) [13]. To identify DEPRGs, we set the values $|\log_{2}FC| > 1$ and P value < 0.05 . We searched for protein-protein interaction networks between these DEPRGs using the STRING database (<https://string-db.org/>) and visualized the resulting network with Cytoscape [14]. For additional analysis, we used the clusterProfiler package (version 4.2.2) to conduct Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses, which provided further insight into the functional significance of the genes [15].

Identification of prognostic platelet-related genes (PRGs)

The initial selection of differentially expressed platelet-related genes (DEPRGs) in the TCGA-PAAD cohort was conducted using univariate Cox regression, with a significance threshold of $P < 0.05$. Subsequently, the Kaplan-Meier algorithm was employed to further refine gene selection, also using a significance level of $P < 0.05$, to ensure the robustness of our findings. Both methodologies identified a common set of prognostic platelet-related genes, demonstrating a robust and consistent association with survival. The 'survival' package (version 3.5-5) was instrumental in facilitating both univariate and multiple Cox regression analyses [16].

Consensus clustering analysis

We employed the Consensus Cluster Plus package (version 1.58.0) to perform consensus clustering on tumor samples from the TCGA-PAAD cohort based on the selected prognostic gene expression profiles [17]. We used the Partitioning Around Medoids (PAM) algorithm and "spearman" as our distance metric, with 500 bootstraps performed on 80% of the train-

Platelet genes predict pancreatic cancer prognosis

ing set patients. We set the clustering number as 2 to 10, and determined the optimal classification by calculating the consensus matrix and cumulative distribution function (CDF). We established the best cluster number based on the cumulative distribution function (CDF). To further validate whether the consistency clustering subtypes were associated with survival, we performed clustering on data from GEO using the same number of clusters following removal of batch effects with the SVA package (version 3.42.0) [18]. We then used the Kaplan-Meier (KM) survival curve to verify whether the subtypes were associated with survival.

Immune infiltration analysis

We analyzed 28 immunocyte characteristic genes extracted from previous research to evaluate the distribution of specific cellular components in the immune microenvironment [19]. The single sample gene set enrichment analysis (ssGSEA) method was used to calculate scores for the 28 immune cells [20]. Differences between subtypes were compared using the Kruskal-Wallis test. In addition, immune cell infiltration was evaluated using the ESTIMATE package (version 1.0.13) [21].

Gene mutation analysis

To compare gene mutation patterns between subtypes, we obtained mutation data from TCGA-PAAD and used the maftools package (version 2.10.15) to generate gene mutation waterfall plots [22]. We also conducted comparative assessments of Homologous Recombination Defects, Fraction Altered, Number of Segments, Nonsilent Mutation Rate, Aneuploidy Score, and Silent Mutation Rate between subtypes.

Establishment and validation of platelet-related risk model

The paper utilized the prognostic genes which had been pre-screened previously. Lasso-Cox regression analysis was then conducted on the TCGA-PAAD cohort utilizing the glmnet package (version 4.1-6) for the purpose of obtaining representative genes and establishing prognostic risk characteristics [23]. The formula for the feature model is Risk Score = $\alpha_i \times \beta_i$ PRGs (α_i represent the regression coefficient of each gene, β_i represent the gene expression value).

Patients were partitioned into cohorts of high and low-risk groups based on the median risk score before using the survminer [24] package (version 0.4.9) to describe and compare the survival curves of both the high and low-risk groups via KM analysis. Moreover, the accuracy of the risk model was evaluated through the assessment of the area under the curve (AUC) of the receiver operating characteristic (ROC) curve, with the “timeROC” package (version 0.4) utilized in R [25]. Finally, the established model was verified with data collected from four cohorts of GEO.

Comparison between risk model and clinical phenotypes

To compare the proportions of various subtypes between different risk groups, `chisq.test` was utilized while `wilcox.test` was used to compare the differences in risk scores among various phenotypes. Also, to illustrate the potent predictive ability of the risk model in terms of predicting survival, we performed univariate and multivariate cox regression analyses, examining the effect of both risk scores and the various clinical phenotypes on survival.

Drug sensitivity of risk model genes

Data including gene expression, drug sensitivity, and cell lines were obtained from the CellMiner [26] database (<https://discover.nci.nih.gov/cellminer/>). Since some missing values (NA) were present in the drug sensitivity data, the `impute.knn` function from version 1.68.0 of the `impute` [27] package was utilized for evaluation and completion of the drug data. Pearson correlation coefficient was then calculated between the gene expression and different drugs within the risk model. The gene and drug screening process involved the use of $P < 0.01$ and $|\text{cor}| > 0.55$ as criteria.

Cell culture and transfection

Pancreatic cancer cell lines HPAC cells were purchased from Cell Bank of Type Culture Collection of Chinese Academy of Sciences. HPAC cells were infected with pLKO.1-shRNA. Silenced cells were selected by puromycin (4 $\mu\text{g}/\text{mL}$) for at least 4 days. All cells were routinely cultured in DMEM/F12 (1:1) + 0.002 mg/ml insulin + 0.005 mg/ml transferrin + 40 ng/ml hydrocortisone + 10 ng/ml epidermal growth

Platelet genes predict pancreatic cancer prognosis

factor + 10% FBS in a humidified incubator containing 5% CO₂ at 37°C.

Western blot

The total cell lines protein was extracted using RIPA Lysis Buffer and PMSF (Thermo Scientific) according to the manufacturer's protocol, and the protein concentration was measured by total protein concentration determination (BCA method). The following primary antibodies: GNA15 polyclonal antibody (Abcam, ab154157, 1:2000) and the GAPDH specific polyclonal antibody (Abcam, ab8245, 1:1000).

Cell proliferation assay

HPAC cells were cultured in 96-well plates (5,000 cells per well). After cell stabilization, Cell Counting Kit-8 reagent (CCK-8) (C0037, Beyotime, China) was added, and the absorbance at 450 nm was detected using a microplate reader at 1, 2, 3, 4, and 5 days to analyze the cell proliferation status.

Cell scratch assay

HPAC Cells were prepared into suspension, add 100 µl of suspension containing 30,000 cells to the scratched insertion well, and remove the insert after the cells overgrow into the insertion well. Then, the cells were cultured in serum free medium and the wound healing status was evaluated at 36 hours.

Transwell assay

5 × 10⁴ cells were plated in the upper chamber in serum-free medium. Filled the culture medium containing fetal bovine serum into the bottom of the well to induce migration and invasion. Cells were incubated for 4-6 hours and stained with hematoxylin and eosin (Boyden chamber). Randomly select five fields and count cells.

Results

Function analysis of differential platelet-related genes (PRGs) in PAAD

We performed a differential analysis of platelet genes, resulting in the identification of 8 down-regulated genes and 202 upregulated genes (Figure S2A and S2B). Subsequently, these 210 differentially expressed genes underwent

enrichment analysis, focusing on the top 10 entries in each category of Gene Ontology (GO) analysis, encompassing molecular function, biological process, and cellular component, as well as the top 30 pathways in the Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis (Figure S2C and S2D). The majority of the enriched results demonstrated a strong association with platelets. Additionally, we constructed a protein-protein interaction network for the genes, revealing prominent interactions of genes such as SRC, MAPK3, RHOA, and HRAS with other genes (Figure S2E).

Selection of prognostic PRGs

We conducted univariate Cox regression analysis on 210 differentially expressed genes and found 74 genes, of which 22 were classified as protective genes and 52 as risk genes (Figure 1A). Subsequently, we employed the Kaplan-Meier algorithm to assess the survival association of the 74 genes identified and ultimately pinpointed 35 genes linked to survival (Figure 1B). Visualization of the univariate Cox regression results for the 74 genes and the Kaplan-Meier curves for the 35 genes was carried out to elucidate their impact on survival.

Clustering analysis of 35 prognostic PRGs

The optimal number of clusters was determined based on the cumulative distribution function (CDF) (Figure 2A). Upon observing the CDF delta area curve, the clustering result was found to be relatively stable when selecting 2 clusters (Figure 2B and 2C). Ultimately, a value of $k = 2$ was chosen, yielding two molecular subtypes: cluster C1 with 73 cases, and cluster C2 with 103 cases. Subsequent analysis of the prognostic characteristics of both molecular subtypes revealed a significant difference in prognosis, with cluster 2 showing poorer prognosis compared to cluster 1 (Figure 2D). Additionally, the clustering results of 35 prognostic genes were validated in data from four GEO datasets, after removing batch effects, demonstrating significant differences in prognosis between the two molecular subtypes (Figure 2E). Furthermore, stromal components, immune components, and ESTIMATE scores in the tumor microenvironment were significantly higher in the C1 group compared to the C2 group (Figure 2F). Moreover, using the ssGSEA method, 28 types of immune cells were scored,

Platelet genes predict pancreatic cancer prognosis

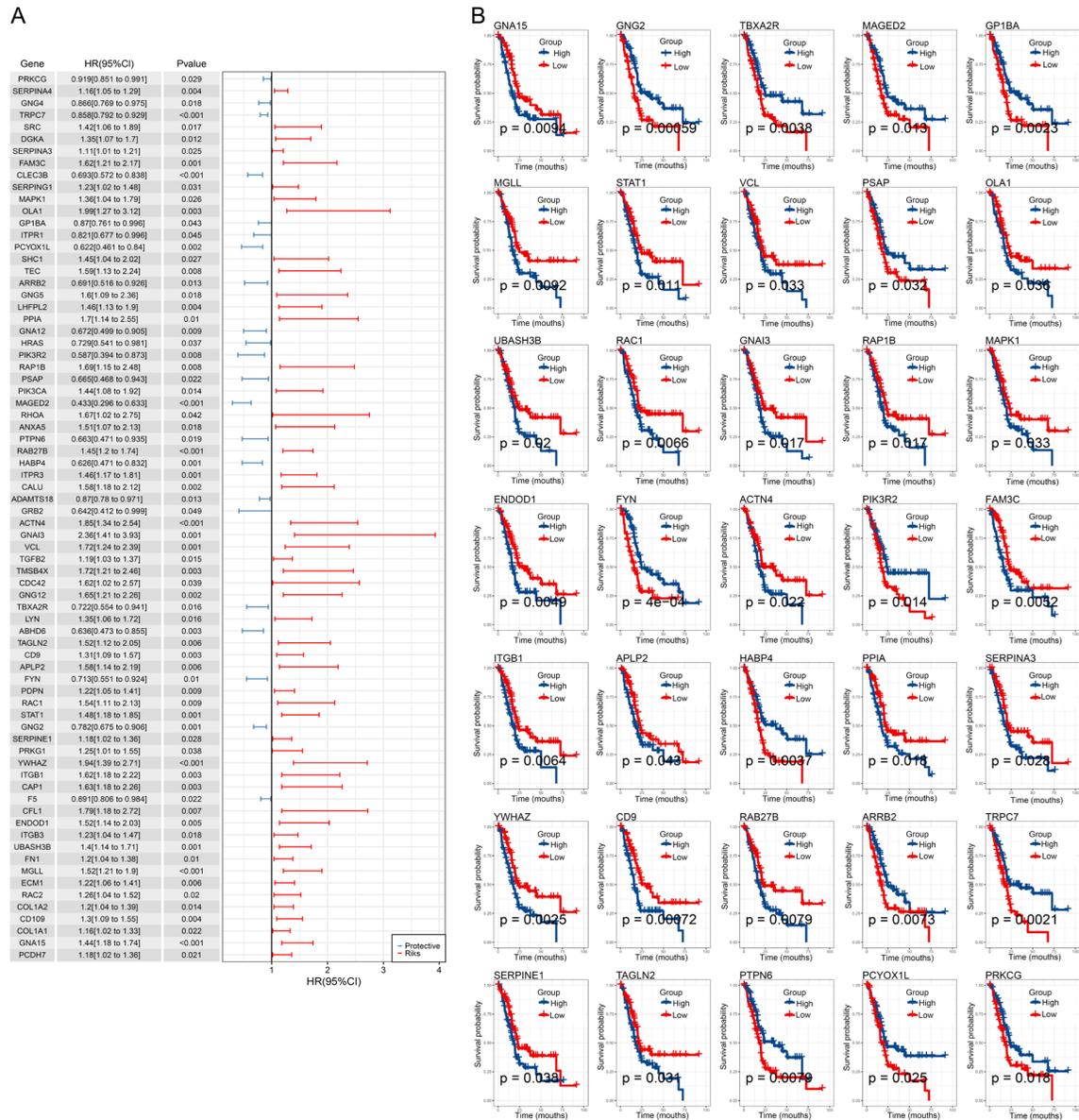


Figure 1. The prognostic significance of 74 platelet genes in PAAD. A: Forest plot displaying 74 platelet genes with $P < 0.05$ in univariate Cox regression analysis. B: Kaplan-Meier (KM) survival curve for 35 platelet genes associated with survival.

and the differences between the two subtypes were tested using the Kruskal-Wallis test, revealing higher scores of activated CD8+ T cells, central memory CD4+ T cells, and natural killer T cells in the C1 group compared to the C2 group (Figure 2G).

Mutation characteristics of the two molecular subtypes

We analyzed the single nucleotide variant (SNV) mutation data from TCGA-PAAD. Figure S3A

illustrates the top 15 genes with the highest mutation frequencies in each subtype. This analysis highlighted a notably higher mutation rate in genes such as KRAS and TP53 in the C2 group. Additionally, we assessed the variance in six key indicators among subtypes: “Homologous Recombination Defects”, “Fraction Altered”, “Number of Segments”, “Nonsilent Mutation Rate”, “Aneuploidy Score”, and “Silent Mutation Rate”. Our findings indicate significant disparities between the C1 and C2 groups in these metrics, with the C1 group

Platelet genes predict pancreatic cancer prognosis

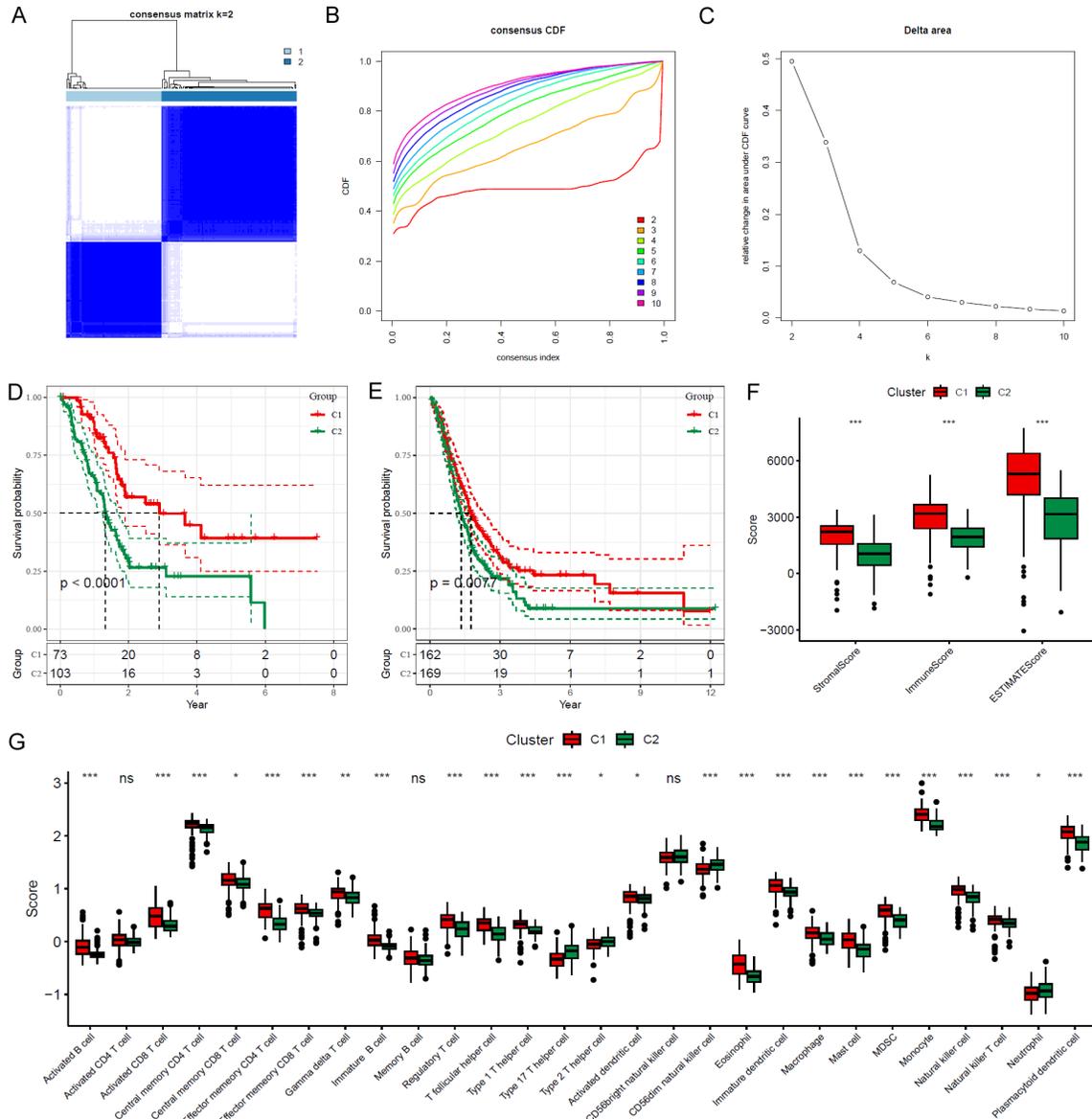


Figure 2. Molecular subtypes based on platelet genes and their characteristics in PAAD. **A:** Heatmap of sample clustering based on consensus $k = 2$. **B:** Cumulative distribution function (CDF) curve of the TCGA sample cohort. **C:** CDF Delta area curve of the TCGA sample cohort, illustrating the clustering number k on the horizontal axis and the relative change in the area under the CDF curve on the vertical axis. **D:** KM survival curve depicting the prognosis of two subtypes in the TCGA-PAAD cohort. **E:** KM survival curve illustrating the prognosis of two subtypes in the GEO cohort. **F:** Differences in ESTIMATE immune infiltration between various molecular subtypes of the TCGA cohort. **G:** Comparative analysis of 28 immune cell scores among different molecular subtypes in the TCGA cohort.

consistently exhibiting lower values than the C2 group, as detailed in [Figure S3B](#).

Establish and evaluation of prognostic risk model

We utilized 35 candidate prognostic-related PRGs to develop a prognostic risk model using

Lasso-Cox regression analysis, as illustrated in [Figure 3A](#) and [3B](#). This model informed the creation of a prognostic scoring system for each of the 18 PRGs, calculated based on gene expression and regression coefficients as follows: Risk Score = $0.069 * GNA15 + 0.091 * MGLL + 0.09 * UBASH3B + 0.214 * SERPINE1 + 0.446 * STAT1 - 0.532 * FYN + 0.095 * CD9 +$

Platelet genes predict pancreatic cancer prognosis

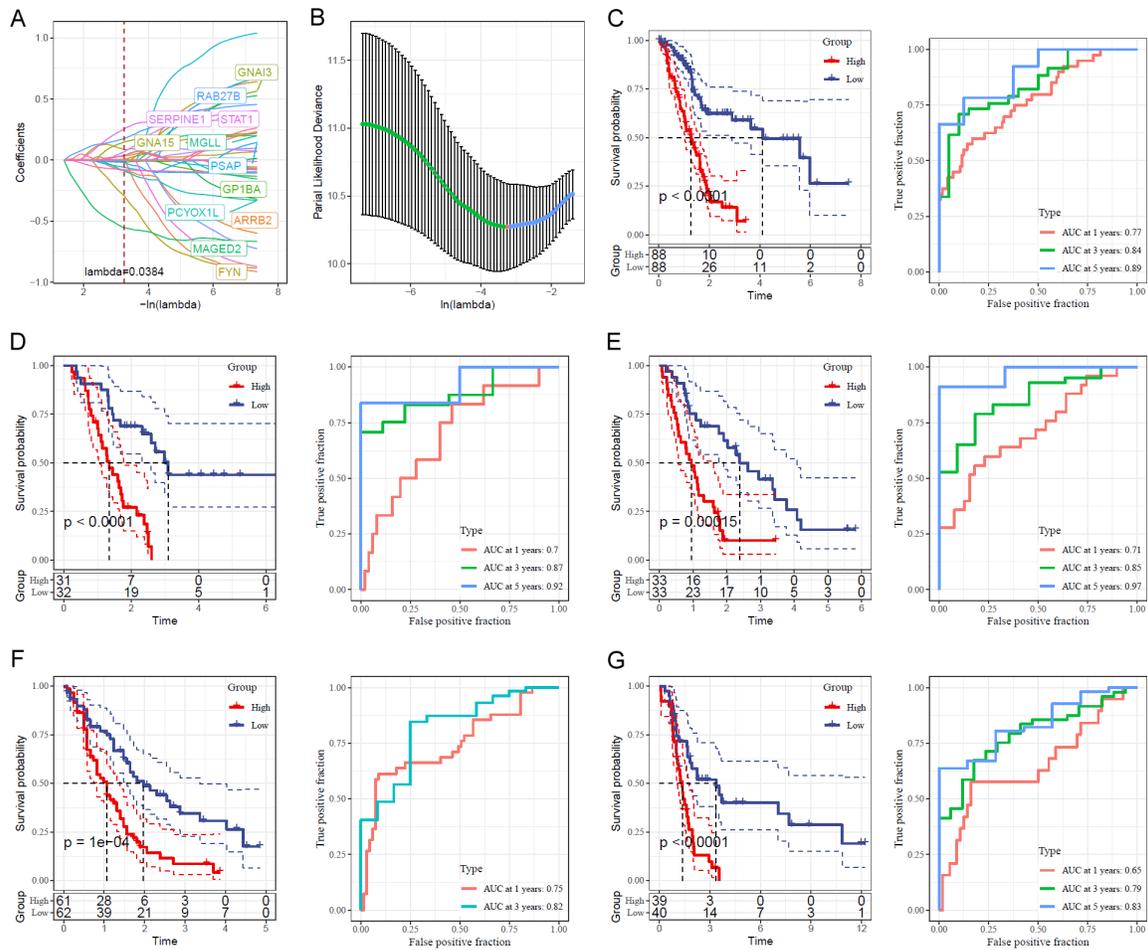


Figure 3. Independent platelet genes detection and risk score establishment in PAAD. A: Trajectory of each independent variable with changing λ values by lasso analysis. B: Confidence interval analysis under varying λ by lasso analysis. C: KM survival curve and ROC curve of risk score, based on the risk model constructed from 18 genes in the TCGA-PAAD cohort. D: KM survival curve and ROC curve of risk score in high and low-risk groups of the GSE57495 cohort. E: KM survival curve and ROC curve of risk score in high and low-risk groups of the GSE62452 cohort. F, G: KM survival curve and ROC curve of risk score in high and low-risk groups of the GSE71729 and GSE85916 cohorts.

0.23 * GNAI3 - 0.103 * HABP4 + 0.281 * RAB27B - 0.711 * MAGED2 - 0.323 * PSAP - 0.031 * ARR2 - 0.271 * PCYOX1L - 0.009 * GP1BA - 0.073 * FAM3C - 0.105 * TRPC7 - 0.106 * PRKCG. We then stratified the TCGA-PAAD cohort patients into high-risk (88 cases) and low-risk (88 cases) groups based on the median risk score. Kaplan-Meier survival analysis revealed that high-risk patients exhibited significantly poorer overall survival compared to their low-risk counterparts, as depicted in **Figure 3C**. This pattern was consistent in external datasets: GSE57495 (**Figure 3D**), GSE62452 (**Figure 3E**), GSE71729 (**Figure 3F**), and GSE85916 (**Figure 3G**), where high-risk patients consistently showed lower overall sur-

vival rates. To assess the prognostic accuracy of our risk features, we calculated the area under the curve (AUC) for 1-year, 2-year, and 5-year receiver operating characteristic (ROC) curves. The AUCs for the TCGA-PAAD cohort were 0.77, 0.84, and 0.89 at 1, 3, and 5 years, respectively. The risk features also demonstrated robust performance in the GEO cohorts, underscoring the reliability and effectiveness of our prognostic assessment.

Comparison of risk score in different clinical-pathological features

To explore the association between Risk Score and tumor clinical characteristics in TCGA-

Platelet genes predict pancreatic cancer prognosis

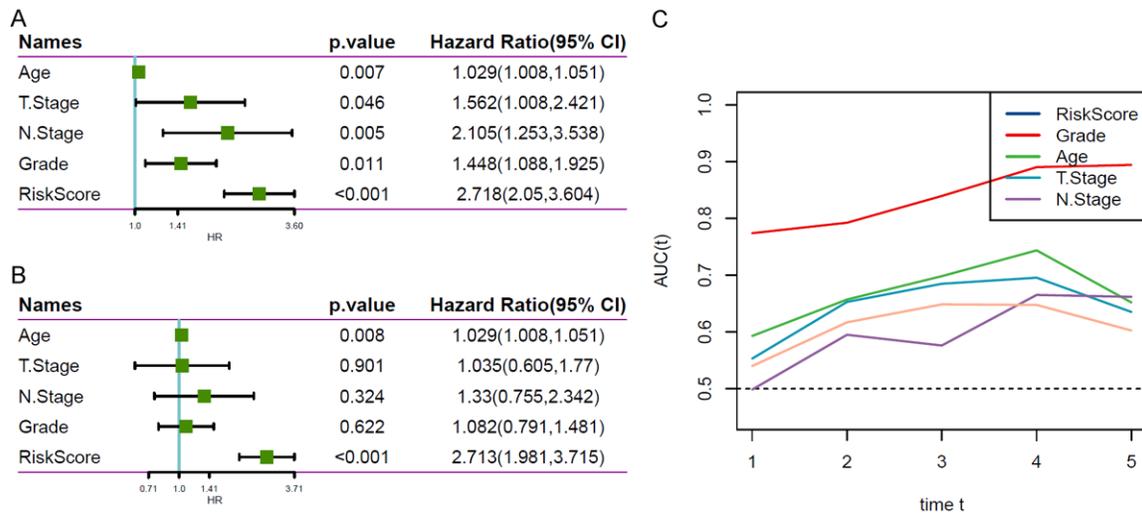


Figure 4. Univariate and multivariate Cox regression analyses of RiskScore with clinical characteristics and AUC time trends. A: Univariate Cox regression analysis of RiskScore and clinical characteristics. B: Multivariate Cox regression analysis of RiskScore and clinical characteristics. C: AUC line chart, with the horizontal axis representing the time in years and the vertical axis indicating the AUC value of the model.

PAAD, we examined the variation of Risk Score across different clinical phenotypes. Our findings indicated a positive correlation between the Risk Score and increased T stage, N stage, and grade (Figure S4A). However, no significant relationship was observed between M stage, age, and Risk Score. This lack of correlation with M stage may be attributed to inadequate follow-up data. Furthermore, we compared the clinical-pathological features across different Risk Score groups within the TCGA cohort, observing analogous trends (Figure S4B).

Comparison of RiskScore with other clinical prognostic variables

We performed both univariate and multivariate Cox regression analyses to further assess the impact of RiskScore and additional clinical variables on survival outcomes. The univariate Cox regression analysis demonstrated that RiskScore presented a notably higher risk compared to other factors ($P < 0.001$), with a hazard ratio (HR) of 2.718 (Figure 4A). Consistently, the multivariate Cox regression analysis yielded similar results ($P < 0.001$, HR = 2.713) (Figure 4B). Additionally, an analysis of the area under the curve (AUC) line graph for these variables revealed that over 1, 2, 3, 4, and 5 years, the AUC for RiskScore consistently exceeded that of other factors, underlining its significance (Figure 4C).

Drug sensitivity of risk model genes

We established a link between three genes - GNA15, HABP4, and ARRB2 - and six drug sensitivities, using a correlation coefficient threshold of $|\text{cor}| > 0.55$ and a P -value < 0.01 as criteria. Our analysis indicated that GNA15 positively correlates with the sensitivities to several drugs, including Methylprednisolone, Chelerythrine, Asparaginase, Fludarabine.1, Fludarabine.2, Zalcitabine, Nelarabine, and Sapacitabine (Figure S5A-H). In addition, HABP4 was positively correlated with the sensitivity to Telatinib (Figure S5I), and ARRB2 showed a similar positive correlation with Sapacitabine's sensitivity (Figure S5J). Remarkably, Sapacitabine's sensitivity was positively associated with both ARRB2 and GNA15 genes (Figure S5H and S5J). Nelarabine exhibited the strongest correlation with the GNA15 gene, with a correlation coefficient of 0.684 (Figure S5G).

Validation the key gene - GNA15 of riskscore in HPAC cells

Initially, we assessed GNA15 expression at the protein level in five pancreatic cancer cell lines using Western blot analysis (Figure 5A). Among these, the HPAC cell line, which exhibited high GNA15 expression, was chosen for knockdown experiments (Figure 5B). Subsequent CCK8

Platelet genes predict pancreatic cancer prognosis

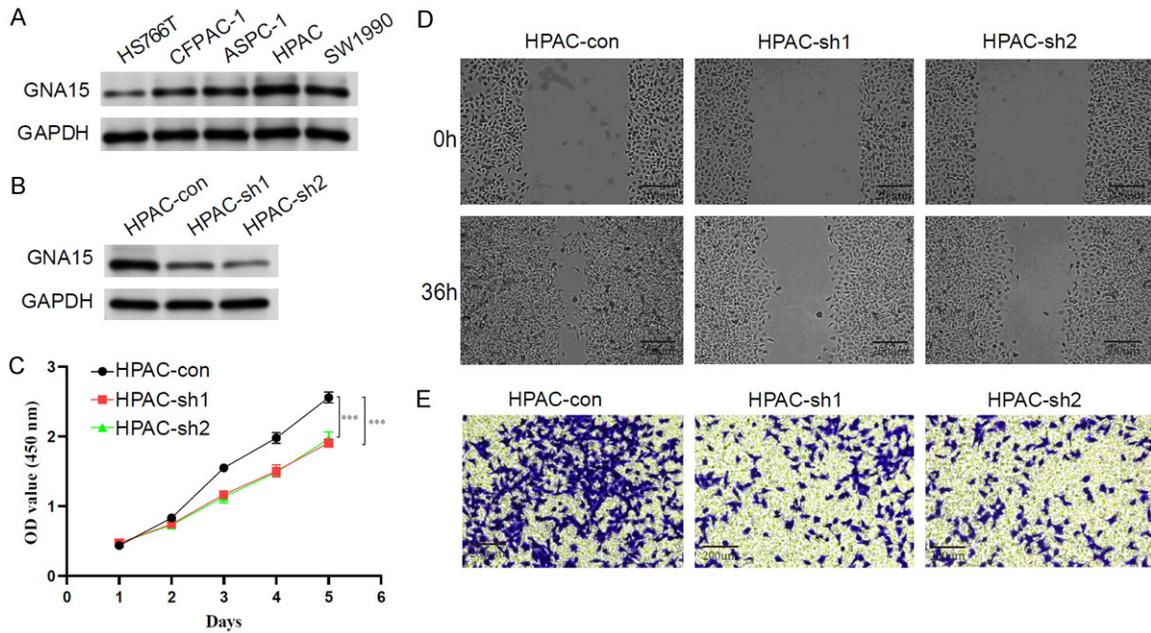


Figure 5. Pancreatic cancer cell expression of GNA15, with knockdown inhibiting proliferation, invasion, and migration. A: GNA15 protein expression levels in HS766T, CFPAC-1, ASPC-1, HPAC, and SW1990 cells. B: Western blot analysis of GNA15 knockdown in HPAC cells. C: CCK8 assay assessing the viability of HPAC-Con, HPAC-sh1, and HPAC-sh2 cells following GNA15 knockdown. D and E: Cell scratch and Transwell assays examining the effects of reduced GNA15 expression on migration and invasion in HPAC-Con, HPAC-sh1, and HPAC-sh2 cells.

assays demonstrated that GNA15 knockdown significantly inhibited the proliferation of HPAC cells (**Figure 5C**). Similarly, cell scratch and transwell assays revealed that GNA15 knockdown markedly reduced the migration (**Figure 5D**) and invasion (**Figure 5E**) capabilities of HPAC cells, respectively.

Discussion

In our study, we identified two molecular subtypes, Cluster C1 and Cluster C2, in pancreatic cancer. Cluster C2 was associated with a poorer prognosis compared to Cluster C1. Notably, the scores for activated CD8⁺ T cells, central memory CD4⁺ T cells, and natural killer T cells were higher in the C1 group than in the C2 group. Additionally, a higher frequency of gene mutations was observed in the C2 group. We also developed and validated a new prognostic prediction model and a platelet-related gene risk score for pancreatic cancer, focusing on the role of platelets. This risk score was found to increase with the progression of T stage, N stage, and grade, and it presented a significantly higher risk than other clinical prognostic factors. It is important to note that the current

5-year survival rate for pancreatic adenocarcinoma (PAAD) is less than 10% [28-30]. The long-term overall survival (OS) rates and prognosis of pancreatic adenocarcinoma (PAAD) remain dismal, largely attributed to its aggressive malignancy, rapid progression, and limited therapeutic alternatives [29, 31]. Platelets are known to facilitate tumor growth and metastasis across various cancers. Studies have indicated that tumor-infiltrating platelets (TIPs) serve as an independent prognostic factor. Incorporating TIPs into the American Joint Committee on Cancer (AJCC) TNM staging system could enhance risk stratification and improve the prediction of surgical outcomes in patients with pancreatic ductal adenocarcinoma (PDAC).

We identified 35 genes related to platelet survival and classified them into two subtypes using CDF analysis. Subsequent prognostic signature analysis revealed significant differences in prognosis between the two molecular subtypes. Moreover, the C1 group exhibited significantly higher tumor microenvironment levels than the C2 group. Further analysis indicated that the C2 group, associated with poor prognosis

Platelet genes predict pancreatic cancer prognosis

sis, showed lower scores for activated CD8+ T cells, central memory CD4+ T cells, and natural killer T cells.

Currently, immune checkpoint inhibitors that target CTLA4 and PD1/PDL1 depend on the activation of effector anti-tumor CD8+ T cells within the tumor microenvironment. Cancers characterized by high tumor mutational burden (TMB) and genomic instability, such as microsatellite instability in colorectal cancer, tend to exhibit increased CD8+ T cell infiltration and are associated with a favorable prognosis. However, pancreatic ductal adenocarcinoma (PDAC) typically presents with a lower TMB and is driven predominantly by chromosomal instability. Most PDAC cases lack the predictive markers for CD8+ T cell infiltration, including high tumor mutation burden and a substantial neoantigen load. Despite this, a minority of PDAC patients with a relatively better prognosis demonstrate significant CD8+ T cell infiltration, suggesting a possible role for these cells in driving anti-tumor immune responses.

CD4+ T cells can be categorized into helper T cells (Th1, Th2, Th17, Th22, Tfh) and regulatory T cells (Treg cells) based on their cytokine secretion profiles. Treg cells, a distinct subset of CD4+ T cells, can suppress the activity of CD8+ T cells and the production of interferon by releasing inhibitory cytokines like IL-10 and TGF- β . This action impedes the activation and proliferation of effector T cells. Additionally, the expression of cytotoxic T-lymphocyte associated antigen 4 (CTLA-4) by Treg cells enhances the IDO pathway in dendritic cells and effector T cells, thereby diminishing the immune function of these effector cells. In pancreatic cancer, an inverse relationship exists between the number of peripheral blood Treg cells and patient survival. In early-stage (I-III) untreated pancreatic cancer, there is nearly a twofold increase in Treg cells in the peripheral blood, constituting about 13% of CD4+ T cells. Patients with a higher proportion of Treg cells within tumor-infiltrating lymphocytes often have a worse prognosis. Accordingly, our findings suggest that the poorer prognosis observed in the C2 group, characterized by platelet-associated genes, might be linked to the tumor immune microenvironment. Moreover, this group exhibited a higher frequency of mutations in genes such as KRAS and TP53.

Utilizing 35 candidate prognostic-related PRGs, we developed a novel prognostic risk model. Patients in the TCGA-PAAD cohort were stratified into two risk groups based on the median risk score, comprising 88 high-risk and 88 low-risk cases. The high-risk group exhibited poorer overall survival compared to the low-risk group. The receiver operating characteristic (ROC) curve analysis for 1, 3, and 5-year intervals showed area under the curve (AUC) values of 0.77, 0.84, and 0.89, respectively. This risk model also demonstrated robust performance across various GEO cohorts, highlighting its accuracy and consistency in prognostic evaluation. Further analysis revealed a correlation between increased T stage, N stage, and grade with higher risk scores. However, no significant association was found between M stage, age, and risk score, potentially due to limited follow-up data for the M stage. We also identified a positive association between three genes (GNA15, HABP4, and ARRB2) and six drug sensitivities, suggesting a potential for predicting drug therapy responses.

Our study, however, has several limitations. Firstly, there is a need to corroborate our current model and risk score using both bulk and single-cell sequencing data. Secondly, further investigations into the underlying mechanisms, both in vivo and in vitro, are warranted. Lastly, additional validation through clinical samples and clinical trials is essential to strengthen the findings.

Conclusion

In conclusion, our study successfully developed a novel prognostic prediction model and risk score for pancreatic cancer, emphasizing the role of platelets. This advancement holds promise for enhancing future drug therapies and improving prognostic predictions in clinical practice.

Acknowledgements

We are grateful to all the participants and the medical staffs in the study.

Disclosure of conflict of interest

None.

Address correspondence to: Hong Yu, Department of General Surgery, Sir Run Run Shaw Hospital,

Platelet genes predict pancreatic cancer prognosis

Zhejiang University School of Medicine, 3th East Qingchun Road, Hangzhou 310016, Zhejiang, P. R. China. E-mail: blueyu000@zju.edu.cn; Zhi-Wei Liu, Department of Blood Transfusion, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, 3th East Qingchun Road, Hangzhou 310016, Zhejiang, P. R. China. E-mail: 3191038@zju.edu.cn

References

- [1] Raimondi S, Maisonneuve P and Lowenfels AB. Epidemiology of pancreatic cancer: an overview. *Nat Rev Gastroenterol Hepatol* 2009; 6: 699-708.
- [2] Siegel RL, Miller KD and Jemal A. Cancer statistics, 2020. *CA Cancer J Clin* 2020; 70: 7-30.
- [3] Gay LJ and Felding-Habermann B. Contribution of platelets to tumour metastasis. *Nat Rev Cancer* 2011; 11: 123-134.
- [4] Takagi S, Sato S, Oh-hara T, Takami M, Koike S, Mishima Y, Hatake K and Fujita N. Platelets promote tumor growth and metastasis via direct interaction between Aggrus/podoplanin and CLEC-2. *PLoS One* 2013; 8: e73609.
- [5] Liu X, Song J, Zhang H, Liu X, Zuo F, Zhao Y, Zhao Y, Yin X, Guo X, Wu X, Zhang H, Xu J, Hu J, Jing J, Ma X and Shi H. Immune checkpoint HLA-E:CD94-NKG2A mediates evasion of circulating tumor cells from NK cell surveillance. *Cancer Cell* 2023; 41: 272-287, e279.
- [6] Wang S, Mo M, Wang J, Sadia S, Shi B, Fu X, Yu L, Tredget EE and Wu Y. Platelet-derived growth factor receptor beta identifies mesenchymal stem cells with enhanced engraftment to tissue injury and pro-angiogenic property. *Cell Mol Life Sci* 2018; 75: 547-561.
- [7] Chen DT, Davis-Yadley AH, Huang PY, Husain K, Centeno BA, Permuth-Wey J, Pimiento JM and Malafa M. Prognostic fifteen-gene signature for early stage pancreatic ductal adenocarcinoma. *PLoS One* 2015; 10: e0133562.
- [8] Yang S, He P, Wang J, Schetter A, Tang W, Funamizu N, Yanaga K, Uwagawa T, Satoskar AR, Gaedcke J, Bernhardt M, Ghadimi BM, Gaida MM, Bergmann F, Werner J, Ried T, Hanna N, Alexander HR and Hussain SP. A novel MIF signaling pathway drives the malignant character of pancreatic cancer by targeting NR3C2. *Cancer Res* 2016; 76: 3838-3850.
- [9] Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, Rashid NU, Williams LA, Eaton SC, Chung AH, Smyla JK, Anderson JM, Kim HJ, Bentrem DJ, Talamonti MS, Iacobuzio-Donahue CA, Hollingsworth MA and Yeh JJ. Virtual microdissection identifies distinct tumor and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat Genet* 2015; 47: 1168-1178.
- [10] Colaprico A, Silva TC, Olsen C, Garofano L, Cava C, Garolini D, Sabedot TS, Malta TM, Pagnotta SM, Castiglioni I, Ceccarelli M, Bontempi G and Noushmehr H. TCGAbiolinks: an R/Bioconductor package for integrative analysis of TCGA data. *Nucleic Acids Res* 2016; 44: e71.
- [11] Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, Porta-Pardo E, Gao GF, Plaisier CL, Eddy JA, Ziv E, Culhane AC, Paull EO, Sivakumar IKA, Gentles AJ, Malhotra R, Farshidfar F, Colaprico A, Parker JS, Mose LE, Vo NS, Liu J, Liu Y, Rader J, Dhankani V, Reynolds SM, Bowlby R, Califano A, Cherniack AD, Anastassiou D, Bedognetti D, Mokraby Y, Newman AM, Rao A, Chen K, Krasnitz A, Hu H, Malta TM, Noushmehr H, Peadarallu CS, Bullman S, Ojesina AI, Lamb A, Zhou W, Shen H, Choueiri TK, Weinstein JN, Guinney J, Saltz J, Holt RA, Rabkin CS; Cancer Genome Atlas Research Network, Lazar AJ, Serody JS, Demicco EG, Disis ML, Vincent BG and Shmulevich I. The immune landscape of cancer. *Immunity* 2018; 48: 812-830, e814.
- [12] Li X, Zhao K, Lu Y, Wang J and Yao W. Genetic analysis of platelet-related genes in hepatocellular carcinoma reveals a novel prognostic signature and determines PRKCD as the potential molecular bridge. *Biol Proced Online* 2022; 24: 22.
- [13] Ritchie ME, Phipson B, Wu D, Hu Y, Law CW, Shi W and Smyth GK. Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* 2015; 43: e47.
- [14] Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B and Ideker T. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res* 2003; 13: 2498-2504.
- [15] Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, Feng T, Zhou L, Tang W, Zhan L, Fu X, Liu S, Bo X and Yu G. clusterprofiler 4.0: a universal enrichment tool for interpreting omics data. *Innovation (Camb)* 2021; 2: 100141.
- [16] Ruiz M, Goldblatt P, Morrison J, Kukla L, Svanacara J, Riitta-Jarvelin M, Taanila A, Saurel-Cubizolles MJ, Lioret S, Bakoula C, Veltsista A, Porta D, Forastiere F, van Eijdsden M, Vrijkotte TG, Eggesbo M, White RA, Barros H, Correia S, Vrijheid M, Torrent M, Rebagliato M, Larranaga I, Ludvigsson J, Olsen Faresjo A, Hryhorczuk D, Antipkin Y, Marmot M and Pikhart H. Mother's education and the risk of preterm and small for gestational age birth: a DRIVERS meta-analysis of 12 European cohorts. *J Epidemiol Community Health* 2015; 69: 826-833.
- [17] Wilkerson MD and Hayes DN. ConsensusClusterPlus: a class discovery tool with confidence

Platelet genes predict pancreatic cancer prognosis

- assessments and item tracking. *Bioinformatics* 2010; 26: 1572-1573.
- [18] Leek JT, Johnson WE, Parker HS, Jaffe AE and Storey JD. The sva package for removing batch effects and other unwanted variation in high-throughput experiments. *Bioinformatics* 2012; 28: 882-883.
- [19] Charoentong P, Finotello F, Angelova M, Mayer C, Efremova M, Rieder D, Hackl H and Trajanoski Z. Pan-cancer immunogenomic analyses reveal genotype-immunophenotype relationships and predictors of response to checkpoint blockade. *Cell Rep* 2017; 18: 248-262.
- [20] Barbie DA, Tamayo P, Boehm JS, Kim SY, Moody SE, Dunn IF, Schinzel AC, Sandy P, Meylan E, Scholl C, Frohling S, Chan EM, Sos ML, Michel K, Mermel C, Silver SJ, Weir BA, Reiling JH, Sheng Q, Gupta PB, Wadlow RC, Le H, Hoesersch S, Wittner BS, Ramaswamy S, Livingston DM, Sabatini DM, Meyerson M, Thomas RK, Lander ES, Mesirov JP, Root DE, Gilliland DG, Jacks T and Hahn WC. Systematic RNA interference reveals that oncogenic KRAS-driven cancers require TBK1. *Nature* 2009; 462: 108-112.
- [21] Yoshihara K, Shahmoradgoli M, Martinez E, Vegesna R, Kim H, Torres-Garcia W, Trevino V, Shen H, Laird PW, Levine DA, Carter SL, Getz G, Stemke-Hale K, Mills GB and Verhaak RG. Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat Commun* 2013; 4: 2612.
- [22] Mayakonda A, Lin DC, Assenov Y, Plass C and Koeffler HP. Maftools: efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* 2018; 28: 1747-1756.
- [23] Friedman J, Hastie T and Tibshirani R. Regularization paths for generalized linear models via coordinate descent. *J Stat Softw* 2010; 33: 1-22.
- [24] Katz J, Lee AC, Kozuki N, Lawn JE, Cousens S, Blencowe H, Ezzati M, Bhutta ZA, Marchant T, Willey BA, Adair L, Barros F, Baqui AH, Christian P, Fawzi W, Gonzalez R, Humphrey J, Huybregts L, Kolsteren P, Mongkolchat A, Mullany LC, Ndyomugenyi R, Nien JK, Osrin D, Roberfroid D, Sania A, Schmiedel C, Silveira MF, Tielsch J, Vaidya A, Velaphi SC, Victora CG, Watson-Jones D and Black RE; CHERG Small-for-Gestational-Age-Preterm Birth Working Group. Mortality risk in preterm and small-for-gestational-age infants in low-income and middle-income countries: a pooled country analysis. *Lancet* 2013; 382: 417-425.
- [25] Blanche P, Dartigues JF and Jacqmin-Gadda H. Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks. *Stat Med* 2013; 32: 5381-5397.
- [26] Luna A, Elloumi F, Varma S, Wang Y, Rajapakse VN, Aladjem MI, Robert J, Sander C, Pommier Y and Reinhold WC. CellMiner cross-database (CellMinerCDB) version 1.2: exploration of patient-derived cancer cell line pharmacogenomics. *Nucleic Acids Res* 2021; 49: D1083-D1093.
- [27] Negro M, Giardina S, Marzani B and Marzatico F. Branched-chain amino acid supplementation does not enhance athletic performance but affects muscle recovery and the immune system. *J Sports Med Phys Fitness* 2008; 48: 347-351.
- [28] Sung H, Ferlay J, Siegel RL, Laversanne M, Soerjomataram I, Jemal A and Bray F. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2021; 71: 209-249.
- [29] Siegel RL, Miller KD, Fuchs HE and Jemal A. Cancer statistics, 2021. *CA Cancer J Clin* 2021; 71: 7-33.
- [30] Miller KD, Fidler-Benaoudia M, Keegan TH, Hipp HS, Jemal A and Siegel RL. Cancer statistics for adolescents and young adults, 2020. *CA Cancer J Clin* 2020; 70: 443-459.
- [31] Zhang SR, Yao L, Wang WQ, Xu JZ, Xu HX, Jin W, Gao HL, Wu CT, Qi ZH, Li H, Li S, Ni QX, Yu XJ, Fu DL and Liu L. Tumor-infiltrating platelets predict postsurgical survival in patients with pancreatic ductal adenocarcinoma. *Ann Surg Oncol* 2018; 25: 3984-3993.

Platelet genes predict pancreatic cancer prognosis

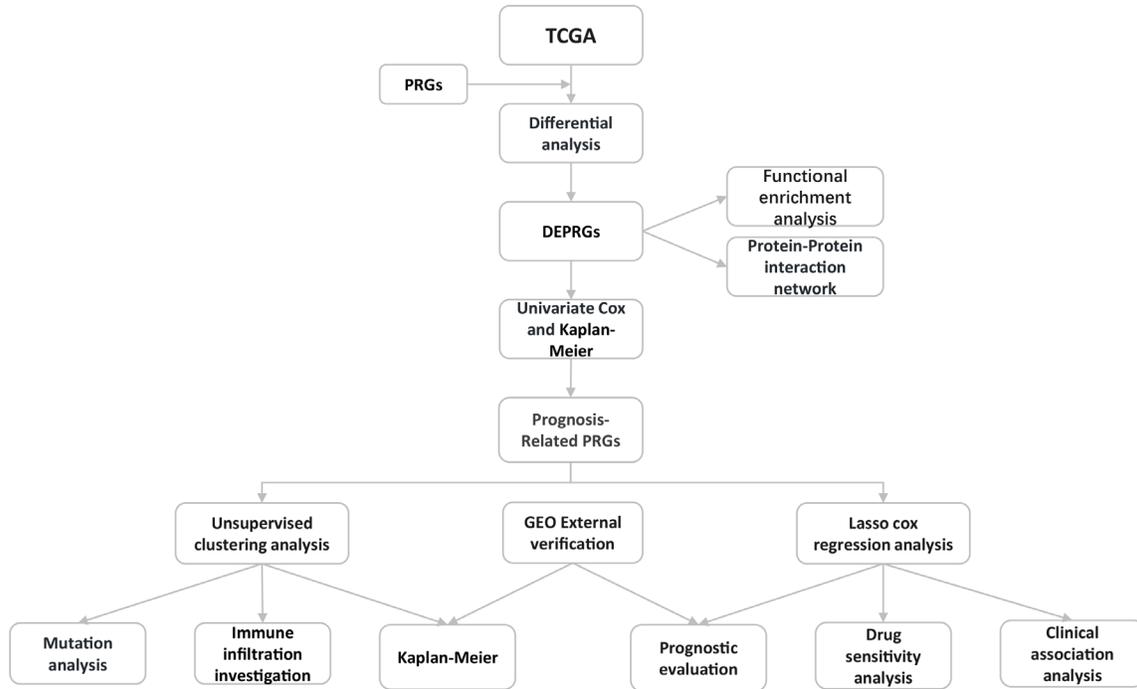


Figure S1. Study Flowchart.

Platelet genes predict pancreatic cancer prognosis

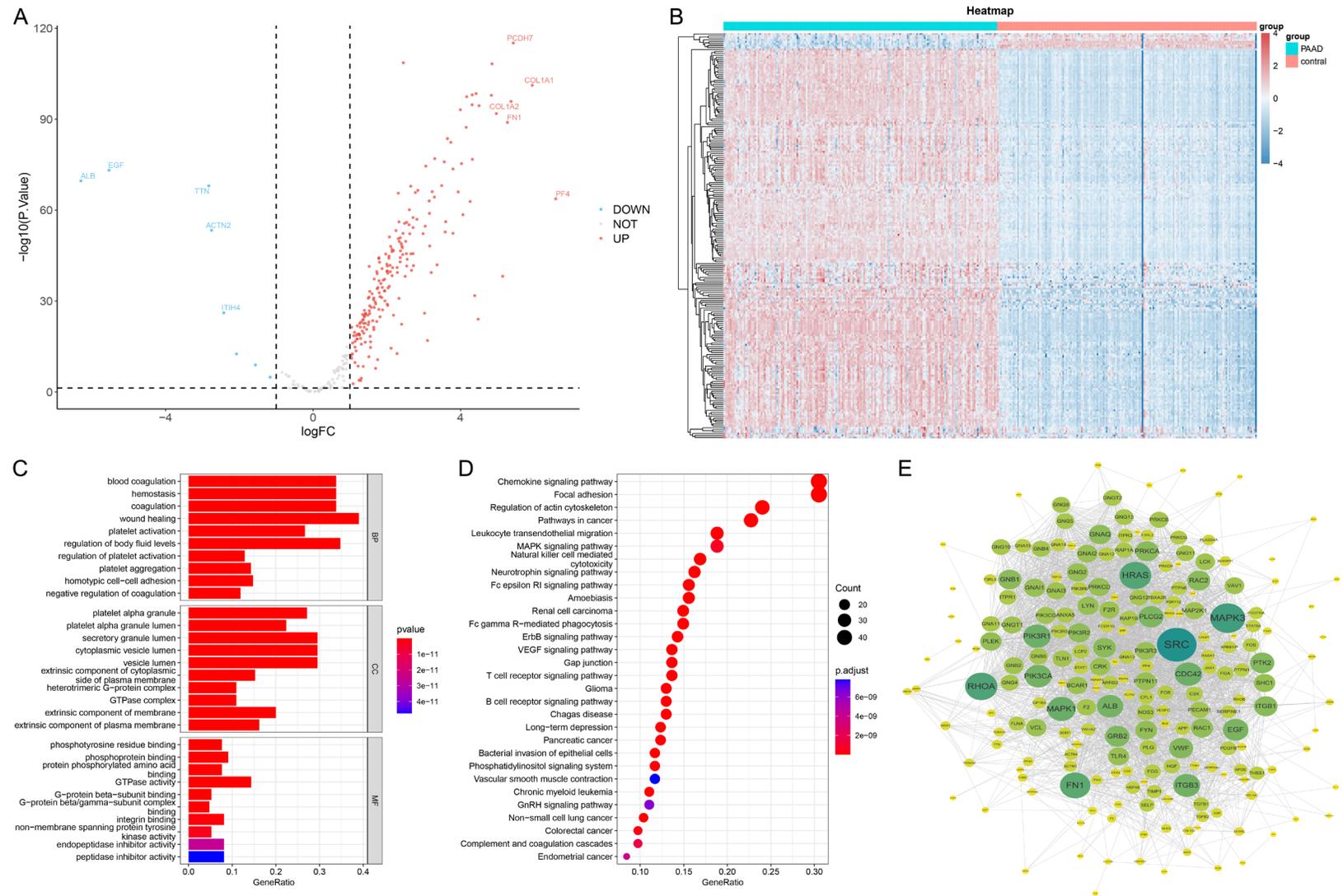


Figure S2. Analysis of differential platelet-related genes (PRGs) and their functions. A: Volcano plot illustrating the differential analysis of platelet genes. B: Heatmap displaying variations in platelet genes. C: Bar graph presenting Gene Ontology (GO) enrichment analysis of differential platelet genes. D: Bubble plot depicting Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of differential platelet genes. E: Protein-protein interaction network of differential platelet genes.

Platelet genes predict pancreatic cancer prognosis

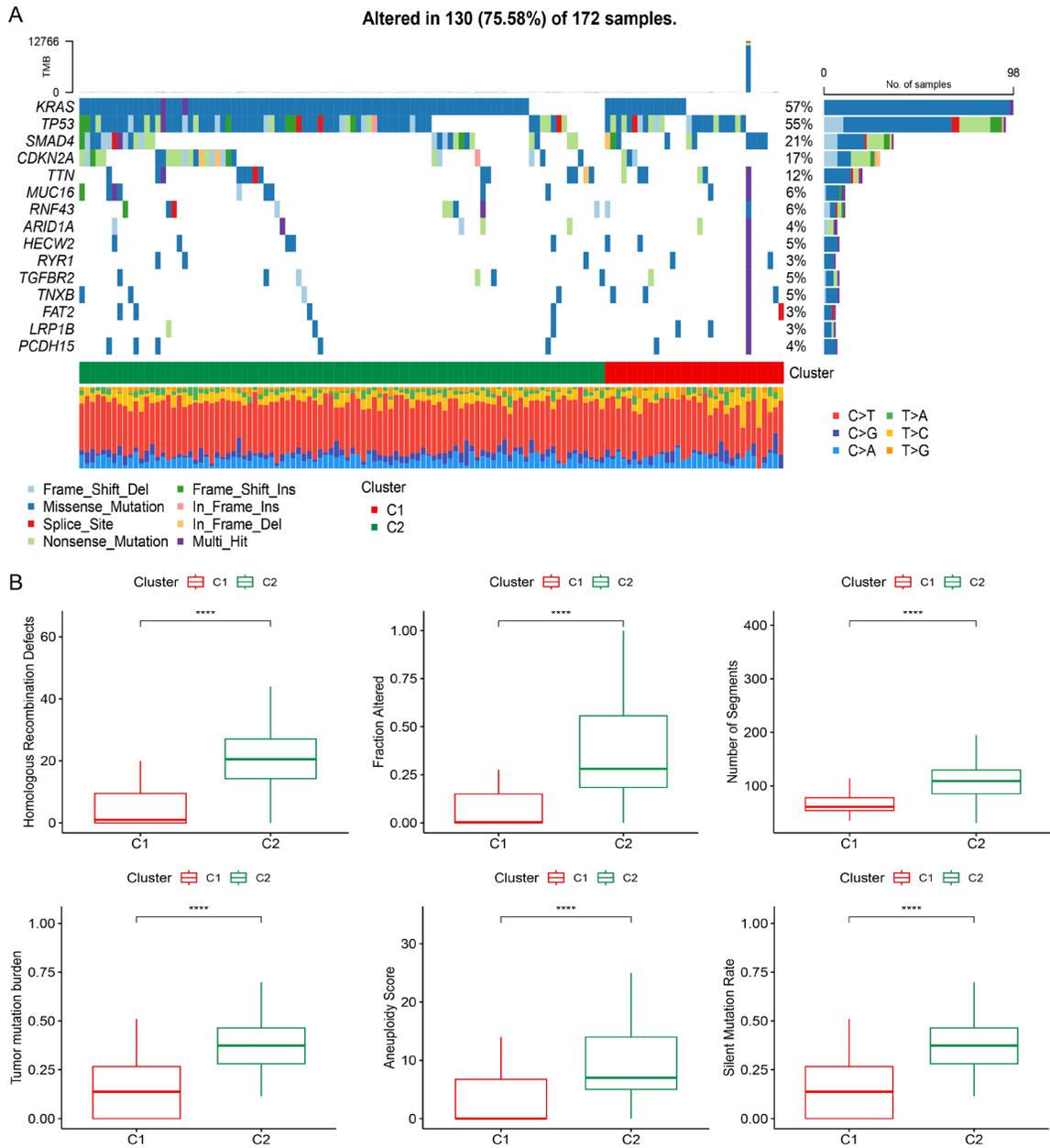


Figure S3. Genomic alterations in TCGA-PAAD cohort molecular subtypes. A: Analysis of somatic mutations in various molecular subtypes of the TCGA-PAAD cohort. B: Comparative analysis of “Homologous Recombination Defects”, “Fraction Altered”, “Number of Segments”, “Nonsilent Mutation Rate”, “Aneuploidy Score”, and “Silent Mutation Rate” across different molecular subtypes of the TCGA cohort.

Platelet genes predict pancreatic cancer prognosis

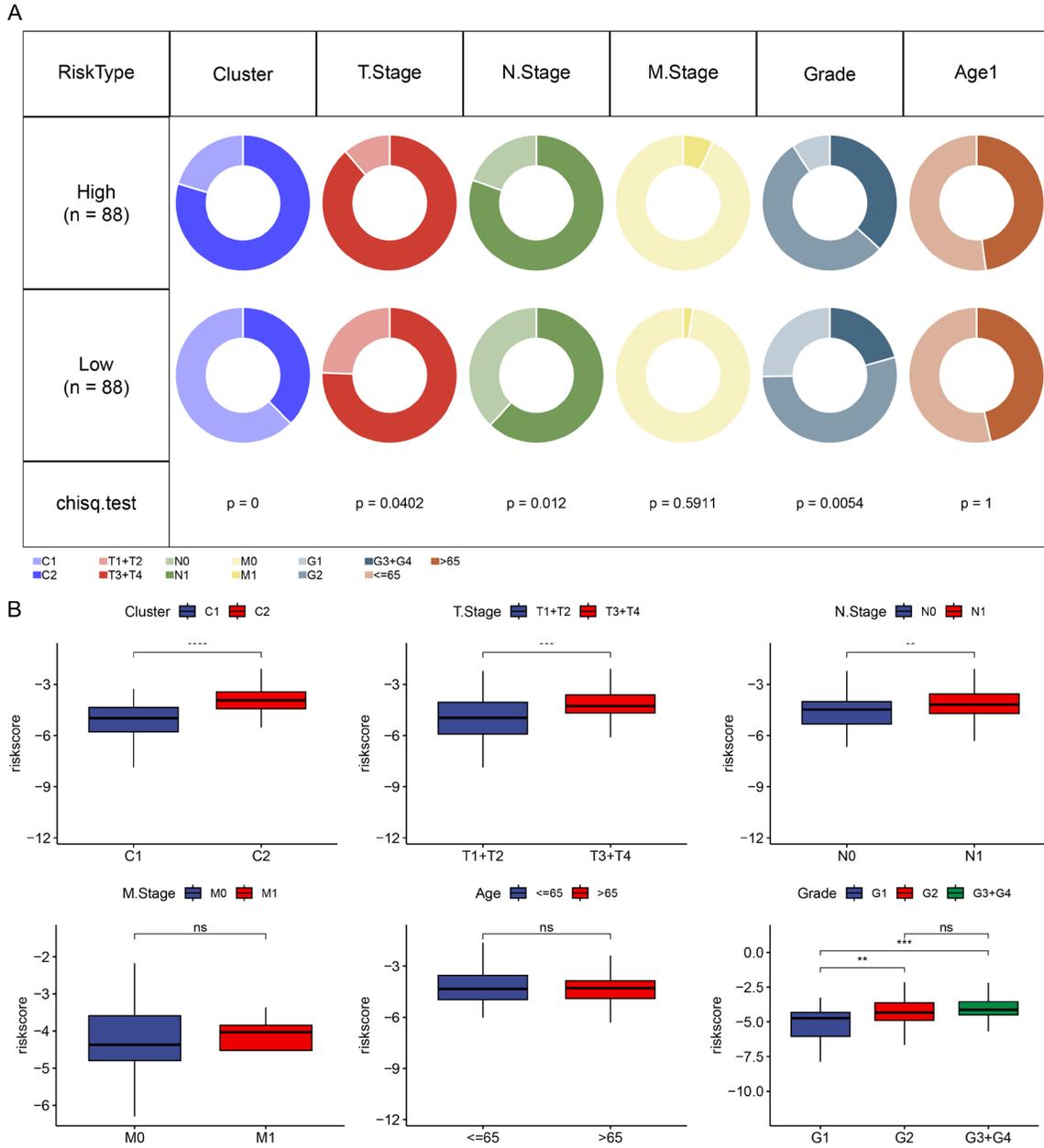


Figure S4. Clinical phenotype comparisons in TCGA cohort RiskScore groups. A: Comparison of clinical phenotypes between RiskScore groups in the TCGA cohort. B: Analysis of RiskScore variations among different phenotypes in the TCGA cohort (Wilcoxon test, *P < 0.05; **P < 0.01; ***P < 0.001; ****P < 0.0001).

Platelet genes predict pancreatic cancer prognosis

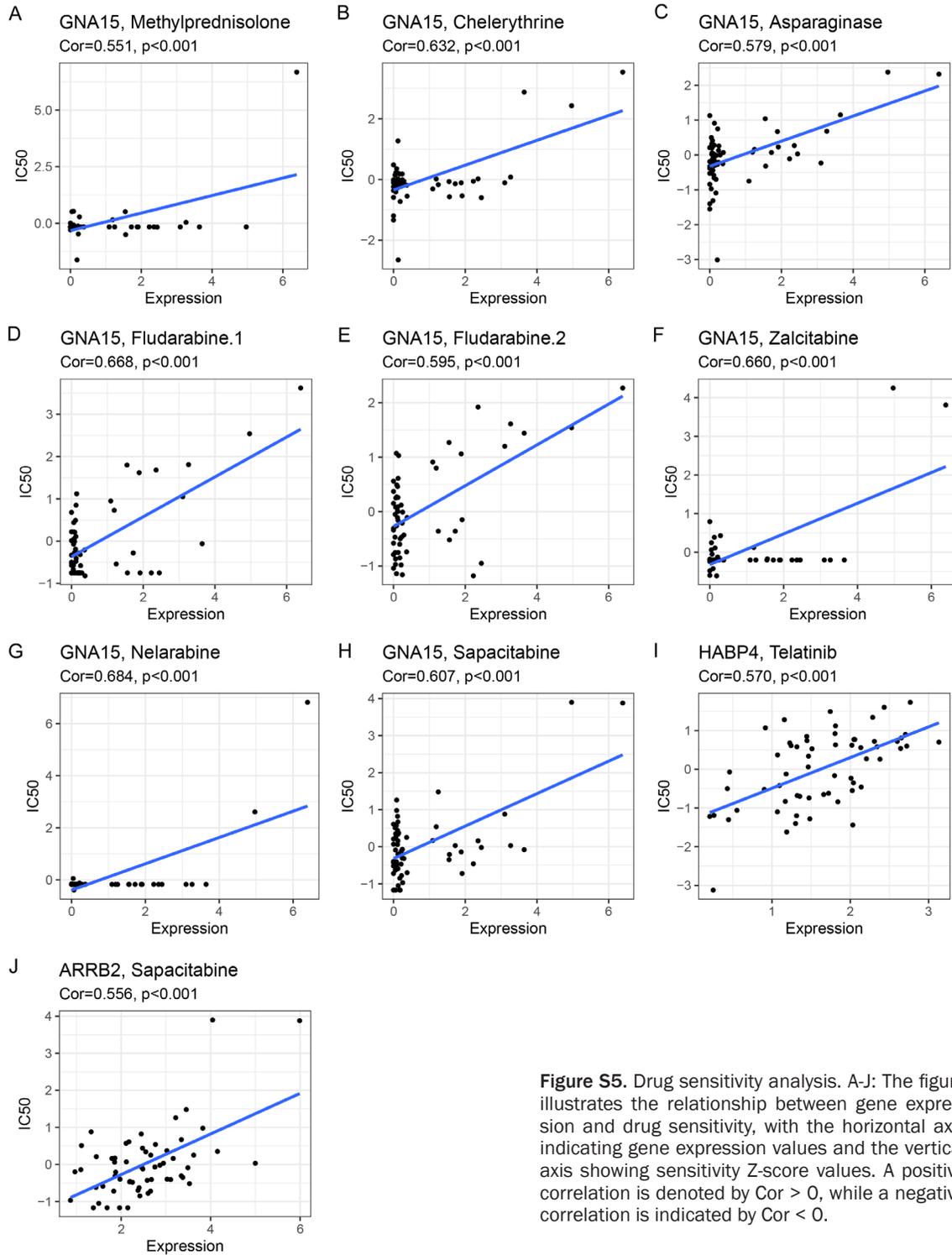


Figure S5. Drug sensitivity analysis. A-J: The figure illustrates the relationship between gene expression and drug sensitivity, with the horizontal axis indicating gene expression values and the vertical axis showing sensitivity Z-score values. A positive correlation is denoted by $Cor > 0$, while a negative correlation is indicated by $Cor < 0$.