PLPP2 as a metabolic and immune marker for predicting survival and enhancing response to anti-PD1 therapy in pancreatic cancer

Guannan Sheng
Tianjin First Center Hospital
Xiaoyan Du
Tianjin Medical University General Hospital Airport Hospital, Tianjin Medical University
Bo Ni
National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin Medical University
Ziyun Liu
National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin Medical University
Chenyan Wu
Beijing Anhua Jinhe Technology Co., Ltd
Kewei Meng
Tianjin First Center Hospital
Yongjie Xie
National Clinical Research Center for Cancer, Tianjin Medical University Cancer Institute and Hospital, Tianjin Medical University
Tao Yang
wayneyt@139.com
Tianjin First Center Hospital

Research Article

Keywords: immune and metabolic signature, PLPP2, pancreatic cancer, immune microenvironment, metabolic

Posted Date: February 14th, 2024

DOI: https://doi.org/10.21203/rs.3.rs-3952664/v1

Additional Declarations: No competing interests reported.

Abstract

Objective

The aim of this study was to establish genetic markers based on metabolic, stromal, and immune factors by analyzing pancreatic cancer (PC) transcriptome datasets to predict prognosis and response to PD-1 therapy in patients with PC.

Methods

We used the pancreatic cancer data set from the TCGA database to identify metabolic-related genetic markers through statistical analysis of artificial intelligence technology. The association between these markers and overall survival (OS) in PC patients was then analyzed. Metabolism, stroma, and immunity were evaluated using GSEA and EPIC algorithms. Finally, external validation was performed on the GEO data set.

Results

PLPP2 was found to be associated with PC metabolism and can effectively predict OS and disease-free survival. Internal verification confirms the accuracy of the mark. PLPP2 was also found to be involved in the metabolism of tumor cells and to regulate the immune system. PLPP2 was evaluated based on clinical relevance, metabolic relevance, immune landscape, and immune checkpoint therapy potential. In vivo experiments showed the potential of PLPP2 as a marker for predicting metabolic status, immune landscape, and response to immune checkpoint inhibitors in PC patients.

Conclusion

PLPP2 is a newly identified marker that predicts stromal, metabolic, and immune features in PC. These findings have potential applications in therapeutic strategies, particularly in the context of immune checkpoint blocking. This study provides crucial insights into the molecular mechanisms of PC, genetic markers that predict prognosis and treatment response, and guides personalized treatment and improved patient outcomes.

Introduction

Pancreatic cancer (PC) is a malignancy characterized by its propensity for early metastasis, accounting for more than 90% of the incidence of various types of pancreatic malignancies[1]. It exhibits notable resistance to established treatment modalities including surgery, chemotherapy, radiotherapy, and molecular targeted therapy[2, 3]. The malignant advancement of PC, spanning from precancerous lesions, such as pancreatic intraepithelial neoplasia (PanINs), to the development of malignancy,

infiltration, and metastasis, is accompanied by the activation of numerous proto-oncogenes and the deactivation of tumor suppressor genes. Studies have found that KRAS, p16/CDKN2A, TP53, DPC4/SMAD4 have high mutation frequency in the process of PC tumorigenesis, resulting in their activation of tumor-promoting function or inactivation of tumor-suppressing function, so they are considered to be the four major driving mutations in the development of PC. In 95% of PC cases, the activation mutation of KRAS occurs at the PanIN stage I, succeeded by the inactivation of the functional tumor suppressor gene p16 (CDKN2A, over 90%). Inactivation of TP53 (about 75 %) and PC4 (SMAD4, about 55 %) freqently occur in PanIN stage III[4]. Metabolic reprogramming within the tumor microenvironment stands as a crucial factor contributing to the unfavorable prognosis observed in PC patients[5]. Collagen ranks as the predominant extracellular matrix component within the tumor microenvironment of PC[6]. PC cells can internalize cleaved collagen fragments or collagen-derived proline, either in a dependent or independent manner[7]. The liberated amino acids resulting from the degradation of engulfed collagen fragments within the lysosome enter the TCA cycle and undergo metabolism, furnishing essential materials for the sustenance of PC cell survival(7). Under conditions of nutrient scarcity, the stimulation and activation of the epidermal growth factor receptor (EGFR) - Pak pathway intensifies the involvement of pinocytosis within PC[8].

The PC microenvironment encompasses a diverse array of interstitial cells. Among these, pancreatic stellate cells (PSC) stand out as specialized fibroblasts unique to the pancreas. Intriguingly, a reciprocal regulatory relationship has been observed between PSC and PC cells. Activated PSC within the PC context release leukemia inhibitory factor (LIF), fostering the progression of PC cells via paracrine signaling[9]. The abnormality of blood vessels in tumor tissue causes insufficient blood supply, which leads to the shortage of nutrients in the tumor microenvironment. Tumor related PSCs secrete alanine (Ala) to provide nutrients for PC cells and activate the TCA cycle, reducing the dependence of tumor cells on serum nutrients such as glucose[10–12].

Immune cells within the PC microenvironment exert a crucial influence on the initiation, progression, therapeutic response, and prognosis of PC. Research has uncovered heightened glycolytic activity in tumor-associated macrophages (TAMs), a phenomenon that contributes to the advancement of PC[13]. In addition, the invasion and metastasis of PC can be significantly eliminated by enhancing the immunosuppressive effect of CD8 + T cells and destroying PI3Kγ, a key lipid kinase in macrophages[14]. Remarkably, upon the knockout of HIF1α in pancreatic tissue, a distinct subgroup of B cells was attracted to the tumor microenvironment, significantly accelerating the advancement of PanIN lesions propelled by KRAS mutations[15].

PC cells undergo metabolic reprogramming to adjust to challenges such as energy and nutrient scarcity, as well as atypical oxidative stress, within the tumor microenvironment. This dynamic adaptation presents novel avenues and tactics for targeted interventions and therapies in pancreatic cancer, grounded in its metabolic attributes. Accordingly, PC cells can be classified into three distinct subtypes based on their metabolic profiles: slow proliferation, glycolysis, and lipogenesis[16]. The level of amino acid and carbohydrate in slowly proliferating PC cells was low, and the cell proliferation was slow. In

glycolytic PC cells, the levels of various intermediate metabolites in the glycolytic pathway and its pathways, such as phosphoenolpyruvate (PEP), glyceraldehyde 3-phosphate, lactic acid and serine, increased, and the expression of genes related to the corresponding glycolytic pathway and pathways, such as PPP pathway, increased. It should be noted that stable isotope tracing experiments found that glycolytic PC cells mainly use the carbon source of glucose to meet the needs of glycolysis, The TCA cycle is maintained by ingesting the carbon source of glutamine[17]. In our study, we constructed a risk model combined with immune status, stromal status and metabolic change which could predict the risk ratio of pancreatic cancer patients. Finally, PLPP2 was selected a targeted one which was inversely correlated with the survival probability of PC patients; and our research provided a new insight on estimating the microenvironment status of PC.

Materials and Methods

The samples for single cell RNA sequencing in our center

Thirteen resected samples from PC patients who received the Whipple-procedure operation were performed the single cell RNA sequencing and subsequent analysis. The specimen was collected following the patient's informed consent and received ethical committee approval. The detailed information of the acquisition of single cell suspension was listed below: sample collection: First, it is recommended that sterile sample handling be used, including the use of nuclease-free reagents and consumables. In order to reduce damage to cells, transfer and centrifugation should be kept to a minimum. At a given centrifugal speed, time and temperature, the cell concentration and size directly affect the efficiency of preparation. Tightly packed cell deposits may require additional manipulation, but may therefore damage cells through shear effects. Of course, centrifugal conditions need to be adjusted at this time. In addition, the use of a sufficiently sized vessel during cell cleaning and resuspension avoids cell agglomeration and agglomeration at high concentrations. Single cell suspension preparation: Entity tissues require mechanical separation or enzymatic digestion at the outset. First, the tissue needs to be physically cut or bladed, and then the cells are isolated by enzymatic digestion. Specific tissue digestive enzymes and digestion time need to be carefully selected. The sample preparation process may cause changes in gene expression patterns in cells, particularly the activation of some stress-related genes. In addition, some sensitive cell subtypes may be harmed in this process. So, the shorter the sample preparation process, the better. On the contrary, if the digestion time is too short, it may lead to incomplete cell separation, and these closely linked cells need to be excluded during subsequent single-cell analysis. The addition of DNase I during cell dissociation prevents the formation of cell aggregates. The single-cell RNA sequencing data produced by our institution has been archived in the Genome Sequence Archive for Human (GSA-Human) under the accession number HRA000433.

The samples for bulk RNA sequencing in our center

We performed bulk RNA sequencing and subsequent analysis on ten resected samples from PC patients who underwent the Whipple procedure. The samples were collected with the informed consent of the

patients and received approval from the ethics committee.

Data acquisition in public database

We enrolled two distinct cohorts of PC patients, each with comprehensive genetic expression profiles and clinical data: The Cancer Genome Atlas (TCGA) cohort, accessible at https://portal.gdc.cancer.gov/, and the GSE62452 cohort from the Gene Expression Omnibus (GEO) database, available at https://www.ncbi.nlm.nih.gov/geo/. All cohorts provide detailed clinical and pathological information, accessible through the GEO database. For our project, we utilized single-cell RNA sequencing data with the accession number GSA: CRA001160, which has been submitted to the Genome Sequence Archive under the project PRJCA001063. Our study incorporated metabolism-related gene sets derived from established gene sets within the Gene Set Enrichment Analysis (GSEA) database. We meticulously adhered to the access regulations of publicly accessible databases, and as the data were sourced from a public database, local ethics committee approval was not required.

Energy Metabolic Molecular Subtypes

In our study, we conducted a comprehensive analysis of PC metabolism-related molecular subtypes utilizing a set of 594 genes. To effectively group all PC samples within the GSE62452 dataset, we employed NMF consensus clustering technique from the R package "NMF". This approach enabled us to obtain robust clustering results. Subsequently, we performed survival analysis and independence tests on the clustered subtypes. To gain further insights, we compared the GSVA scores among different groups. Additionally, we utilized the TIMER tool to derive immune scores for the identified subtypes. This multifaceted analysis provided us with a comprehensive understanding of the metabolic and immune characteristics of the different PC subtypes.

DEG Identification and Bioinformatics Analysis

We employed the "DESeq2" R package to calculate the differentially expressed genes (DEGs) within the identified subtypes. Specifically, we focused on DEGs with false discovery rate (FDR) values below 0.05 and an absolute log2-fold change exceeding 1. Subsequently, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) functional enrichment analyses using these selected DEGs. These analyses allowed us to elucidate the biological pathways and molecular processes associated with the observed transcriptional differences between the subtypes

Risk Model Construction and Validation

We employed the expression data of the identified DEGs from the training cohort to construct a risk score model. Our initial step involved assessing the influence of each DEG on the overall survival (OS) of PC patients. And we employed the univariate Cox proportional risk regression model, utilizing a significance threshold of log-rank $P < 0.01$. Subsequently, we utilized LASSO-Cox regression to streamline the selection of genes in our model, effectively reducing their number. The risk score model was constructed by incorporating individual normalized gene expression values adjusted by their respective LASSO-Cox coefficients. To ensure the robustness of our model, we subjected it to validation using both internal and

external validation cohorts. To calculate the risk score of each sample, we applied our formulated equation, and we visually represented the distribution of risk scores using the "timeROC" R package. By leveraging the Gordon index, we determined a cutoff value to classify samples into high- and low-risk groups. We employed a log-rank test to compare survival differences between these two groups, and their overall survival (OS) was visually depicted through Kaplan-Meier (KM) survival curves. This comprehensive methodology facilitated the development of a reliable risk score model for predicting patient outcomes based on the gene expression profile.

Prognostic value of the risk signature in training and validation group

Patients were stratified into two distinct groups, namely high- and low-risk, using the median value of the risk score. To evaluate the prognostic capacity of the risk signature, we generated Kaplan-Meier (K-M) survival curves employing the Log-rank test. Additionally, we harnessed the "survival ROC" R package to compute the area under the curve (AUC) values for the receiver operating characteristic (ROC) curves corresponding to 1-, 2-, 3-, and 5-year survival periods. These analyses provided a comprehensive insight into the predictive performance of the risk score model across different time frames. These evaluations were conducted to assess the performance of the two signatures. To avoid redundancy, different sentence structures and words were used, and some keywords were replaced.

Gene set enrichment analysis (GSEA)

To elucidate the intrinsic molecular mechanisms linked to the high- and low-risk clusters, we conducted Gene Set Enrichment Analysis (GSEA). For this purpose, we employed the reference gene set h.all.v7.2.symbols.gmt (Hallmarks) extracted from the Molecular Signature Database. GSEA serves as a potent analytical technique, facilitating the identification of biologically significant gene sets that exhibit enrichment within a designated phenotype or sample group. This approach aids in unraveling the functional pathways and biological processes that underlie the observed differences between the high- and low-risk categories. By comparing the high- and low-risk groups, GSEA can reveal the potential molecular pathways that contribute to the different clinical outcomes observed between the two groups. The results of the GSEA analysis will provide insights into the biological processes that are driving the differential prognosis in pancreatic cancer.

Analysis of Immune Infiltration Characteristics in Relation to Metabolism-Related Risk Score Subgroups in PC

We investigated the immune infiltration patterns in pancreatic cancer using the RNA-seq data from the GSE62452 database (GEO) and explored the potential associations between immune cells and our prognostic signature consisting of six genes. To analyze the immune infiltration characteristics, we utilized seven online tools, including CIBERSORT and TIMER. We then compared the relative proportions of immune cells in the six genes that constituted our prognostic signature, providing further insight into the potential role of immune cells in pancreatic cancer prognosis. By investigating the immune infiltration

patterns and potential associations with our prognostic signature, we aimed to uncover potential therapeutic targets and improve the clinical management of pancreatic cancer.

Cell culture

The KPC cell line was cultivated in RPMI 1640 cell culture medium supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin, within a specialized cell incubator maintained at 37°C under a controlled atmosphere of 2% carbon dioxide. The construction of KPC-PLPP2-sh cell line was based on NM_001302389.

Subcutaneous tumor model

C57BL/6J mice, aged between 4 and 6 weeks, were procured from GemPharmatech. All mice were anesthetized with isoflurane and lower abdominal hair was cleared. Disinfect the injection site with 75% alcohol before inoculation. The needle was inserted from the left inguinal region of the mouse and punctured forward about 1cm for subcutaneous injection. All animal experiments were conducted in accordance with protocols that received approval from the Ethics Committee and Institutional Review Board of Tianjin Medical University Cancer Institute & Hospital.

Flow cytometry

Fresh tumor tissue was mechanically homogenized to obtain a single-cell suspension. This suspension was subsequently centrifuged at 500 g for 10 minutes, after which the supernatant was removed. The resulting pellet was then resuspended in a phosphate buffer solution. 200ul single cell suspension was drawn into a flow tube, and flow antibodies such as CD45, CD3, CD8, PD-1 were added, and incubated on ice for 30–60 minutes in the dark. Granzyme B and TNFα were labeled according to the protocol by BD.

Statistical analysis

In this study, statistical analyses and graphical representations were carried out using R (version 3.6.3) or GraphPad Prism (version 8.3.0). Continuous variables were analyzed using the t-test, while Fisher's exact test or the chi-square test was used for categorical variable comparisons. To assess disparities among Kaplan-Meier (K-M) survival curves, the log-rank test was applied. A significance threshold of less than 0.05 (two-tailed) was considered indicative of statistical significance. These rigorous analyses upheld the credibility and robustness of our findings.

Results

Identification and analysis of metabolism and immune-related differentially expressed genes in PC

Pancreatic cancer cells are prone to construct a metabolically vigorous microenvironment for providing an enriched nutritional requirement. Besides, the formation of immune suppressive environment is positively correlated with the development of PC. For further exploring an immune and metabolism-

related gene set, we made a systematic and precise screened strategy combined with both immune and metabolism factors **(Figure. 1A)**. First, RNA bulk sequencing was conducted using resected tissue samples obtained from ten patients diagnosed with PC. These patients were closely monitored through the utilization of ultrasonography **(Figure. 1B)**. The staining intensity of CD8 + T cells was assessed using immunohistochemistry (IHC), and immune infiltration analysis was conducted using the CIBERSORT software. These methods were employed to enhance the quantification of CD8 + T cells within archived tissues. Subsequently, a selection of 1242 Differentially Expressed Genes (DEGs) was made based on the "L_CD8 + T cell" group. Following this, a separate cohort of thirteen PC patients who had undergone resected Whipple surgery was subjected to single-cell RNA sequencing **(Figure. 1C)**. Subsequently, we performed reduction and cell cluster identification analysis with the expression matrix. Combined with the stained intensity of CD8 + T cell, six PC patients were divided into “L_CD8 + T cell” and “H_CD8 + T cell” groups; and then 1132 DEGs were selected from “L_CD8 + T cell” group comparing with “H_CD8 + T cell” group. Finally, we then performed differential gene analysis (DEGs) according to the proportion of CAF in ten PC patients with EPIC software, and 972 differential genes were obtained in the high fibrosis group **(Figure. 1D)**. Totally, 34 intersected genes in “L_CD8 + T cell’ group between the single cell RNA sequencing and bulk RNA sequencing were identified. Finally, we collected 1385 metabolism-related genes (protein metabolism, sugar metabolism, lipid metabolism) in MSigDB database to perform the Sperman’s correlation analysis with 254 intersected genes, and 54 targeted genes were identified **(Figure. 1E)**. Subsequently, we conducted Gene Set Enrichment Analysis (GSEA) based on the Kyoto Encyclopedia of Genes and Genomes (KEGG) to delve into the potential functions of these genes and gain insights into the modified metabolic landscape within PC. The outcomes of functional enrichment analysis using KEGG indicated a significant correlation between metabolic differentially expressed genes (mDEGs) and various pathways. Notably, these pathways encompassed Carbon metabolism, Glycerophospholipid metabolism, Biosynthesis of amino acids, and Spinocerebellar ataxia pathways, as depicted in Fig. 1G. Furthermore, the 5 most enriched Gene Ontology (GO) terms associated with metabolic mDEGs included fatty acid metabolic process, small molecule catabolic process, glycerol-lipid metabolic process, phospholipid metabolic process, and organic acid biosynthetic process. These findings are visually depicted in Fig. 1F.

Correlation analysis between clinicopathological characteristics and the metabolism-related gene signature

To construct a robust and comprehensive metabolism-related risk model (MRM) for prognostic prediction, we initiated the process by identifying 222 common genes from the pool of metabolic differentially expressed genes mDEGs. Subsequently, we employed LASSO Cox regression analysis to scrutinize the key mDEGs linked to DFS. This approach ensured the inclusion of crucial genes while minimizing potential overfitting and optimizing the predictive capacity of the model. Through cross-validation, we identified six genes (MGLL, MET, MBOAT2, PLPP2, DDIT4, B4GALT2) that demonstrated the lowest partial likelihood deviance **(supplementary Figure. 1A-B)**. Moreover, an array of correlation analyses was executed to explore the relationship between the metabolism-related gene signature and various

clinicopathological features. Notably, the DFS and overall survival (OS) rates were notably lower in cases exhibiting high expression of these metabolism-related genes, as opposed to those with low expression. This observation strongly suggested that elevated expression of these markers was indicative of a heightened likelihood of recurrence **(supplementary Figure. 2)**. For further validating the function of these metabolism-related genes in different subgroups, K-M survival analysis in subgroups of clinical characteristics (T stage, histological grade, pathologic stage, N stage) indicated that the selected metabolic and immune-related genes were closely and inversely correlated with OS **(Figure. 2 and supplementary Figure. 3)**.

Construction of a metabolism-related risk model to predict the overall survival and disease-free survival of PC patients

Following the computation of the MRS using the developed model, we classified PC patients into two distinct subgroups: high- and low-risk categories. This categorization was achieved by utilizing the median value of the MRS as a threshold to demarcate the two subgroups **(Figure. 3A)**. Our analysis unveiled a significant finding: patients exhibiting higher MRS exhibited notably lower DFS and OS rates in comparison to those with lower MRS values. This compelling evidence indicates that MRS holds the potential to serve as a valuable prognostic factor for PC **(Figure. 3B-C)**. We also examined the expression levels of the six genes in the model across the two risk subgroups. Furthermore, we found that MRS was positively associated with several clinicopathological features, including T stage, histological grade, pathologic stage, residual tumor, and N stage, which was demonstrated using a nomogram **(Figure. 3D)**. ROC curve analysis suggested that MRS could effectively predict patient survival. The Area Under the Curve (AUC) values for predicting 1-, 3-, and 5-year overall survival (OS) were 0.686, 0.726, and 0.799, respectively **(Figure. 3E).** Moreover, multi-variate cox analysis identified MRS as an independent prognostic factor for pancreatic cancer **(Figure. 3F)**.

Validation of the prognostic value of metabolism-related risk model in internal and external cohorts

To verify the predictive value of this model, we performed internal validation with partial data from dataset GSE62452. According to **Figure. 4A**, the efficacy of our Metabolism-Related Risk Model (MRM) in prognosticating the outcomes of PC patients was substantiated through meticulous validation using both internal and external validation datasets. Our findings consistently demonstrated that individuals with elevated MRS exhibited decreased DFS and OS durations, heightened susceptibility to recurrence, and overall poorer prognosis. These conclusions were corroborated by the results of Kaplan-Meier survival analysis **(Figure. 4B)**. Subgroup analysis reinforced the observation of a strong and adverse correlation between the risk score and OS. In other words, higher risk scores were consistently associated with diminished OS across different clinical subgroups **(Figure. 4C-D)**. To enhance the precision of prognosticating PC patient outcomes, we formulated a nomogram integrating variables linked to OS. This nomogram exhibited commendable performance, as evidenced by the calibration curve. This tool offers an effective means to quantitatively estimate individual patient prognoses, facilitating more accurate

clinical predictions **(Figure. 4E-F)**. In the validation cohort of GSE62452, the AUC values for predicting 1-, 2-, 3-, and 5-year BCR (Biochemical Recurrence) were 0.634, 0.750, and 0.939, respectively. These impressive AUC values signify that the MRM holds the potential to serve as a robust and precise clinical tool for forecasting DFS outcomes in patients with PC **(Figure. 4G)**.

Furthermore, the outcomes of KEGG enrichment analysis unveiled a significant correlation between upregulated DEGs in the high-risk group and several pathways, including Carbon metabolism, Glycerophospholipid metabolism, Biosynthesis of amino acids, and Spinocerebellar ataxia pathways. Equally noteworthy, the top five enriched GO terms for these DEGs encompassed fatty acid metabolic process, small molecule catabolic process, glycerol-lipid metabolic process, phospholipid metabolic process, and organic acid biosynthetic process **(Figure. 4H-I)**.

Identification of PLPP2 as a key regulator in the immune suppressive environment of PC

WGCNA, which stands for weighted gene co-expression network analysis, is a method used to analyze gene expression patterns across multiple samples. It facilitates the grouping of genes that exhibit similar expression profiles, enabling the examination of correlations between gene modules and distinct phenotypes or traits. This technique aids in uncovering potentially meaningful biological insights and associations within complex datasets. In this study, WGCNA was used to analyze PC patients in the TCGA database. The clustering tree **(Figure. 5A)** revealed no significant differences between the samples included in the WGCNA analysis. By applying the scale-free topological model and average connectivity, the optimal soft threshold of 8 was selected **(Figure. 5B-C)**. The WGCNA analysis resulted in a gene cluster tree **(Figure. 5D)** where each leaf and branch represented a gene and a co-expression module, respectively. The heat map generated from the WGCNA analysis **(Figure. 5E)** illustrated the correlation between different modules and clinical characteristics. Five modules, including the yellow module, were obtained from the WGCNA analysis, with the yellow module exhibiting the highest correlation with immune suppressive traits (correlation value = -0.33; $p < 0.05$, number of module members = 22). The characteristic key genes in the yellow module were plotted in a scatter plot **(Figure. 5F)**, with PLPP2 being identified as the most important gene for further exploration **(Figure. 5G)**.

The clinical characteristics and molecular biological functions of PLPP2 in PC

In the data set of 178 cases of pancreatic cancer in TCGA database, we confirmed that PLPP2 expression significantly predicted poor survival and prognosis in patients (HR = 1.769, P value = 0.028) **(Figure. 6A-C, Supplementary table1, Supplementary table2))**. In order to further explore the function of PLPP2 in PC, we continued to use the single cell transcriptome dataset in GEO database (GSE12345). In the preceding analysis, we substantiated that ductal cell 2 exhibited traits of a malignant ductal cell type, whereas ductal cell 1 demonstrated characteristics consistent with normal ductal epithelium. The identification of KRT19 and EPCAM confirmed the distinctive features of the ductal epithelial subgroup. Concurrently, the validation of MUC1 and FXYD3 further affirmed the malignant attributes of ductal cells exhibiting

malignancy **(Figure. 6D-E)**. To investigate the subcellular distribution of PLPP2 within pancreatic cancer cells, we established that PLPP2 predominantly localizes within the nucleus **(Figure. 6F)** with the U, M cell lines in the *human protein atalas* website, and the expression of this molecule was also found to be significantly higher in cancer than that in precancerous tissues from the immunohistochemical level **(Figure. 6G)**. We continue to confirm by timeROC analysis that PLPP2 molecules have considerable predictive efficiency for the survival rates of pancreatic cancer :1 (AUC = 0.626), 3 (AUC = 0.647) and 5 -year (AUC = 0.766) **(Figure. 6H)**. Given that the PLPP2 screening process primarily accounted for immune cell infiltration, tumor cell stemness, and the ratio of cancer-associated fibroblasts within the microenvironment, we conducted Spearman's correlation analysis between PLPP2 molecules and established stemness genes (SOX9, POU5F1, NANOG, PROM1), genes linked to Epithelial-Mesenchymal Transition (EMT) (ZEB1, SNAI1, SNAI2, VIM), and genes associated with Activated Pancreatic Stellate Cells (PSCs) (COL1A1, ACTA2, PDGFB, TGFB1). This correlation analysis aimed to uncover potential associations and interactions between PLPP2 and these specific molecular characteristics. A significant positive correlation among them was presented with co-expressed heatmap, and the correlation was statistically significant **(Figure. 6I-K)**. Through the assessment of immune infiltration in patients categorized into high- and low- PLPP2 expression groups, a noteworthy observation emerged. Elevated PLPP2 expression exhibited a notably and significantly negative correlation with the levels of infiltrating T cells, Tgd, Tfh, pDC, NK cells, cytotoxic cells, CD8 + T cells, and B cells. This correlation demonstrated statistical significance with a p-value < 0.05. This finding implies a potential role of PLPP2 in influencing immune cell infiltration within the tumor microenvironment **(Figure. 6L)**. The obtained results strongly imply that PLPP2 exerts a crucial role within the microenvironment, contributing to the advancement of pancreatic cancer. To comprehensively understand its function, our next steps involve delving into in vitro and in vivo investigations. These experiments will shed light on the specific mechanisms through which PLPP2 influences pancreatic cancer development and provide deeper insights into its potential as a therapeutic target.

Expression of PLPP2 predicted the efficacy of anti-PD-1 therapy

Our results indicated that high expression of PLPP2 could result in a significant immune-suppressive microenvironment which may be tightly correlated with metabolic and fibrosis factors. Then we next evaluated the effect of PLPP2 on immune checkpoint inhibitor therapy in public datasets. We found that high expression of PLPP2 could significantly predict the worse response to anti-PD-1 and CTLA-4 therapy comparing with low expression of PLPP2 in Riza cohort 2018(Anti-PD-1/CTLA-4), Gao cohort 2018(Anti-PD-1/CTLA-4), Van cohort 2021(Anti-PD-L1) datasets **(Figure. 7A-C)**. Subsequently, receiver operating characteristic curve (ROC) shown that high expression of PLPP2 predicted more non-responders of immune checkpoint inhibitors in Dizier cohort 2013(AUC = 0.527), Van cohort 2021(AUC = 0.607), Riaz cohort 2018(AUC = 0.650) and Kim cohort 2019(AUC = 0.553) **(Figure. 7D-G)**. And then the effect on drug sensitivity analysis in two obvious datasets with different expression of PLPP2 showed that high expression of PLPP2 indicated the therapy resistance of immune checkpoint inhibitors **(Figure. 7H)**.

Tumoral PLPP2 promoted tumor proliferation and suppressed immune microenvironment in PC

The diversity in the immune microenvironment significantly contributes to chemotherapy resistance, recurrence, and unfavorable prognoses. Recent advances in immunotherapy and associated combination treatments have instilled optimism for patients with advanced tumors. A systematic exploration of tumor immune microenvironment heterogeneity plays a pivotal role in guiding treatment choices, predicting effectiveness, optimizing strategies, and identifying novel immunotherapy targets. We hypothesized that the heightened expression of tumoral PLPP2 triggers a reshaping of the suppressive immune environment, chiefly propelling the proliferation of pancreatic malignant tumors. To confirm these effects, we established stable KPC-PLPP2-scramble/sh cell lines expressing PLPP2. Subsequently, we executed in vivo experiments to examine the correlation between tumoral PLPP2 and CD8 + T cell infiltration. Remarkably, our results echoed the same trend, with decreased tumor burden observed in the KPC-PLPP2-sh group within the subcutaneous C57BL/6 tumor mice model. This signifies a promising direction for further research on the role of PLPP2 in shaping the tumor immune microenvironment and its implications for potential therapeutic interventions **(Figure. 8A-B)**. Furthermore, our inference suggests that one of the primary mechanisms underlying the suppression of tumor proliferation could be linked to the modulation of the immune microenvironment through upregulation of PLPP2. To delve deeper into these immune microenvironment changes, we conducted a thorough investigation into the infiltration and functional alterations of CD8 + T cells within the two groups, utilizing flow cytometry. The preliminary results, as depicted in **Figure. 8(C-D)**, indicated a higher percentage of CD8 + T cells in the PLPP2-sh group. This observation marks a significant stride in unraveling the intricate interplay between PLPP2 and the immune response, providing a promising avenue for unraveling potential therapeutic strategies. Next, we further explored the effect of PLPP2 deficiency on function of CD8 + T cell (TNFα, Granzyme). As shown in **Figure. 8E-F)**, the CD8 + T cell function in PLPP2-sh group was significantly elevated compared with PLPP2-scramble group. Additionally, we extended our investigation to assess the alterations in PD-1 expression levels within CD8 + T cells. The findings revealed that the PLPP2-scramble group exhibited elevated PD-1 expression, a characteristic associated with reduced sensitivity to anti-PD-1 therapy. Conversely, the PLPP2-sh group demonstrated a notable reduction in PD-1 levels within CD8 + T cells, especially when compared to the PLPP2-scramble group. This observation suggests that the deficiency of PLPP2 might render the tumor microenvironment more responsive to anti-PD-1 therapy. This insight could hold significant implications for optimizing treatment strategies and enhancing therapeutic responses **(Figure. 8I-J)**. The success of immunotherapy largely hinges on the interplay between tumor cells and immune regulation within the TME. In this dynamic interplay, the TME can either facilitate or suppress the immune response, thereby playing a pivotal role in determining the effectiveness of immunotherapy. Therefore, comprehending the complex relationship between immunotherapy and TME not only sheds light on the underlying mechanism of action, but also offers valuable insights into developing novel strategies to enhance the therapeutic efficacy of immunotherapy.

Discussion

Cell metabolic reprogramming is a key characteristic of tumor cells and plays a critical role in tumor initiation and progression[18]. Indeed, the changes in intracellular and extracellular metabolites exert a profound impact on various biological aspects, including gene expression, cellular diversity, and the TME[19]. PC undergoes metabolic reprogramming during disease progression, from initial PanIN to advanced PC[20]. Indeed, while PC can exhibit a self-limiting tendency, individuals with high-risk PC encounter notably low survival rates. This underscores the urgency for the development of a reliable prognostic signature, which holds significant potential in assessing and guiding the treatment of this challenging disease. Such a signature could play a crucial role in identifying patients at higher risk and tailoring therapeutic strategies accordingly, ultimately aiming to improve patient outcomes and enhance overall survival.

In our study, we focused on identifying and analyzing metabolic mDEGs in the PC database GSE62452, which were found to be enriched in lipid and amino acid-related metabolic processes. Through consensus clustering analysis based on these mDEGs, we were able to classify patients into two subgroups with significant differences in DFS dindicating that PC metabolism is heterogeneous and that patients with different metabolic patterns have distinct clinical outcomes.

After identifying mDEGs enriched in lipid and amino acid-related metabolic processes in the PC database GSE62452, LASSO Cox regression analysis was used to screen for critical mDEGs associated with survival. Among the six genes identified (MGLL, MET, MBOAT2, PLPP2, DDIT4, B4GALT2), MGLL plays a crucial role in pain and nociception and is linked to cancer tumorigenesis and metastasis[21–23]. the proto-oncogene MET encodes a protein with tyrosine kinase activity and is implicated in various cellular processes, including cell growth, proliferation, and migration. Mutations or aberrant activation of the MET gene have been linked to the development of several cancers, including papillary renal carcinoma. The dysregulation of MET signaling can contribute to tumor progression by promoting cell survival, angiogenesis, and metastasis, making it an important target for cancer research and potential therapeutic interventions[24, 25]. MBOAT2, a major yeast lysophospholipid acyltransferase, has similarities to mammalian membrane-bound O-acyltransferase (MBOAT) proteins. DDIT4 is a mitochondrial and tumor-related protein involved in anti-tumor therapy resistance, proliferation, and invasion[26, 27]. B4GALT2 is one of seven beta-1,4-galactosyltransferase (beta4GalT) genes, encoding type II membrane-bound glycoproteins that appear to have exclusive specificity for the donor substrate UDP-galactose[28–31]. However, the specific mechanisms of these key genes in PC metabolism, immunotherapy, and drug response are yet to be fully understood.

Using the GSE62452 data set as the validation set, further analysis revealed a strong correlation between MRS and the clinical characteristics and prognosis of PC patients, indicating that tumor metabolism changes with disease progression. Our model based on metabolically related genes was found to be a reliable independent prognostic indicator, as confirmed by survival analysis, ROC curve analysis, and univariate multivariate COX regression analysis. The model also exhibited a relatively high accuracy in

predicting tumor prognosis, suggesting its potential as a clinical tool for PC evaluation and treatment. However, one of the limitations of this study is the need for further in-depth research into the specific mechanisms underlying the key genes' role in PC metabolism, immunotherapy, and drug response.

Cell metabolism and TME are closely related, as evidenced by numerous studies[32]. PC can be treated with various methods, including surgery, chemotherapy, radiotherapy, and immunotherapy. Our study aimed to establish a model-based risk score to quantify tumor metabolism and to objectively reveal its correlation with the immune microenvironment, providing guidance for clinical treatment of PC patients. The high-risk group showed an abundance of B cells, CD4 + cells, and CD8 + cells, which can inform treatment decisions. We also investigated the correlation between the six genes in our model and immune checkpoints such as CD44 and CD86. Using the Weighted Gene Co-expression Network Analysis (WGCNA) method, we identified PLPP2 as a potential marker, which is a critical member of the Phospholipid phosphatase (PLPP) family. PLPPs are widely expressed in human tissues and play a role in cell signal transduction, and their upregulation has been observed in various cancers, including pancreatic adenocarcinoma, gliomas, and lung adenocarcinoma[33]. PLPP2, specifically, plays a crucial role in breast cancer development and occurrence[34]. Our analysis revealed that PLPP2 participates in tumor cell metabolic processes and affects the immune status of the TME. We evaluated PLPP2's clinical relevance, metabolic relevance, immune landscape, and its potential for immune checkpoint therapy. Furthermore, we used PLPP2 as a marker to predict the metabolic status, immune landscape, and response to anti-PD-1 therapy in vivo experiments in PC patients. However, we need to conduct further in vitro experiments to explore the specific mechanism of PLPP2.

Declarations

Declaration of conflict of interest: There are no conflicts to declare.

Funding

This study was supported by the Tianjin health science and technology project (ZC20213).

Conflict of interest

The authors have no relevant financial or non-financial interests to disclose.

Contributions

Guannan Sheng, Yongjie Xie and Tao Yang wrote the manuscript text, Xiaoyan Du and Bo Ni drew pictures, Ziyun Liu, Chenyan Wu and Kewei Meng rovided suggestions, and all authors reviewed the manuscript and participated in discussions to form this article.

Data availability

The datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Ethics approval

All animal experimentation procedures adhered to approved protocols by the IACUC and Tianjin Medical University Cancer Institute and Hospital.

References

1. Rhim AD, Mirek ET, Aiello NM, Maitra A, Bailey JM, McAllister F, Reichert M, Beatty GL, Rustgi AK, Vonderheide RH, Leach SD, Stanger BZ. EMT and Dissemination Precede Pancreatic Tumor Formation. Cell 2012; 148:349–361.
2. Kim MP, Gallick GE. Gemcitabine Resistance in Pancreatic Cancer: Picking the Key Players. Clin Cancer Res 2008; 14:1284–1285.
3. Huang L, Jansen L, Balavarca Y, Molina-Montes E, Babaei M, Geest L van der, Lemmens V, Eycken LV, Schutter HD, Johannesen TB, Fristrup CW, Mortensen MB, Primic-Žakelj M, Zadnik V, Becker N, Hackert T, Mägi M, Cassetti T, Sassatelli R, Grützmann R, Merkel S, Gonçalves AF, Bento MJ, Hegyi P, Lakatos G, Szentesi A, Moreau M, Velde T van de, Broeks A, Sant M, Minicozzi P, Mazzaferro V, Real FX, Carrato A, Molero X, Besselink MG, Malats N, Büchler MW, Schrotz-King P, Brenner H. Resection of pancreatic cancer in Europe and USA: an international large-scale study highlighting large variations. Gut 2019; 68:130.
4. Li J-T, Wang Y-P, Yin M, Lei Q-Y. Metabolism remodeling in pancreatic ductal adenocarcinoma. Cell Stress 2019; 3:361.
5. Brunner JS, Finley LWS. Metabolic determinants of tumour initiation. Nat Rev Endocrinol 2023; 19:134–150.
6. Yamauchi M, Barker TH, Gibbons DL, Kurie JM. The fibrotic tumor stroma. J Clin Investig 2018; 128:16–25.
7. Olivares O, Mayers JR, Gouirand V, Torrence ME, Gicquel T, Borge L, Lac S, Roques J, Lavaut M-N, Berthezène P, Rubis M, Secq V, Garcia S, Moutardier V, Lombardo D, Iovanna JL, Tomasini R, Guillaumond F, Heiden MGV, Vasseur S. Collagen-derived proline promotes pancreatic ductal adenocarcinoma cell survival under nutrient limited conditions. Nat Commun 2017; 8:16031.
8. Lee S-W, Zhang Y, Jung M, Cruz N, Alas B, Commisso C. EGFR-Pak Signaling Selectively Regulates Glutamine Deprivation-Induced Macropinocytosis. Dev Cell 2019; 50:381–392.e5.
9. Shi Y, Gao W, Lytle NK, Huang P, Yuan X, Dann AM, Ridinger-Saison M, DelGiorno KE, Antal CE, Liang G, Atkins AR, Erikson G, Sun H, Meisenhelder J, Terenziani E, Woo G, Fang L, Santisakultarm TP, Manor U, Xu R, Becerra CR, Borazanci E, Hoff DDV, Grandgenett PM, Hollingsworth MA, Leblanc M, Umetsu SE, Collisson EA, Scadeng M, Lowy AM, Donahue TR, Reya T, Downes M, Evans RM, Wahl

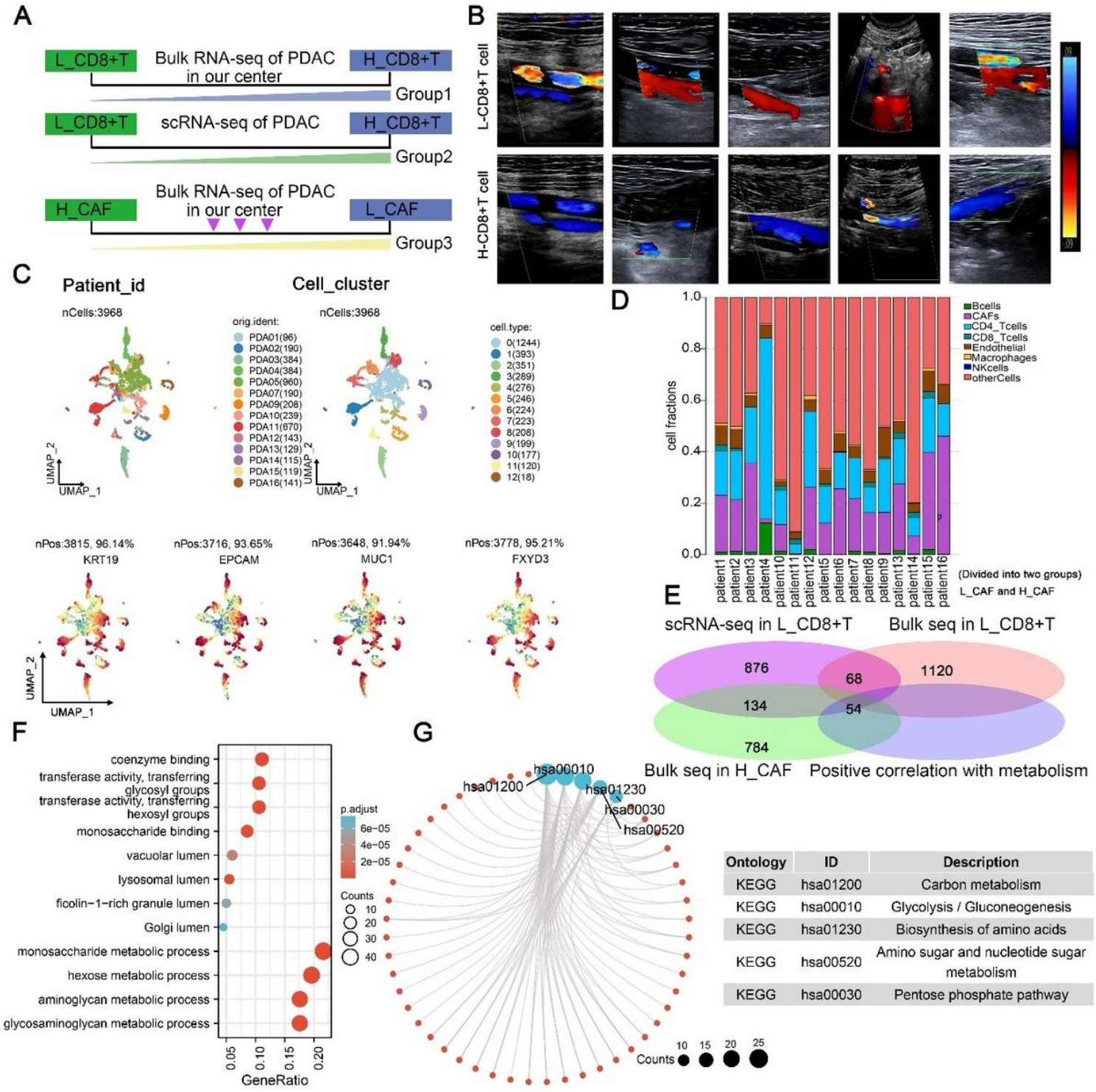
GM, Pawson T, Tian R, Hunter T. Targeting LIF-mediated paracrine interaction for pancreatic cancer therapy and monitoring. Nature 2019; 569:131–135.

10. Sousa CM, Biancur DE, Wang X, Halbrook CJ, Sherman MH, Zhang L, Kremer D, Hwang RF, Witkiewicz AK, Ying H, Asara JM, Evans RM, Cantley LC, Lyssiotis CA, Kimmelman AC. Pancreatic stellate cells support tumour metabolism through autophagic alanine secretion. Nature 2016; 536:479–483.
11. Feig C, Gopinathan A, Neesse A, Chan DS, Cook N, Tuveson DA. The Pancreas Cancer Microenvironment. Clin Cancer Res 2012; 18:4266–4276.
12. Kamphorst JJ, Nofal M, Commisso C, Hackett SR, Lu W, Grabocka E, Heiden MGV, Miller G, Drebin JA, Bar-Sagi D, Thompson CB, Rabinowitz JD. Human Pancreatic Cancer Tumors Are Nutrient Poor and Tumor Cells Actively Scavenge Extracellular Protein. Cancer Res 2015; 75:544–553.
13. Penny HL, Sieow JL, Adriani G, Yeap WH, Ee PSC, Luis BS, Lee B, Lee T, Mak SY, Ho YS, Lam KP, Ong CK, Huang RYJ, Ginhoux F, Rotzschke O, Kamm RD, Wong SC. Warburg metabolism in tumor-conditioned macrophages promotes metastasis in human pancreatic ductal adenocarcinoma. OncoImmunology 2016; 5:e1191731.
14. Kaneda MM, Cappello P, Nguyen AV, Ralainirina N, Hardamon CR, Foubert P, Schmid MC, Sun P, Mose E, Bouvet M, Lowy AM, Valasek MA, Sasik R, Novelli F, Hirsch E, Varner JA. Macrophage PI3Kγ Drives Pancreatic Ductal Adenocarcinoma Progression. Cancer Discov 2016; 6:870–885.
15. Lee KE, Spata M, Bayne LJ, Buza EL, Durham AC, Allman D, Vonderheide RH, Simon MC. Hif1a Deletion Reveals Pro-Neoplastic Function of B Cells in Pancreatic Neoplasia. Cancer Discov 2016; 6:256–269.
16. Daemen A, Peterson D, Sahu N, McCord R, Du X, Liu B, Kowanetz K, Hong R, Moffat J, Gao M, Boudreau A, Mroue R, Corson L, O'Brien T, Qing J, Sampath D, Merchant M, Yauch R, Manning G, Settleman J, Hatzivassiliou G, Evangelista M. Metabolite profiling stratifies pancreatic ductal adenocarcinomas into subtypes with distinct sensitivities to metabolic inhibitors. Proc Natl Acad Sci 2015; 112:E4410–E4417.
17. Li J-T, Lei M-Z, Lei Q-Y, Yin M. [A Review of Metabolic Stress and Development of Pancreatic Cancer]. Sichuan Da Xue Xue Bao Yi Xue Ban J Sichuan Univ Méd Sci Ed 2021; 52:5–10.
18. Pavlova NN, Thompson CB. The Emerging Hallmarks of Cancer Metabolism. Cell Metab 2016; 23:27–47.
19. Boroughs LK, DeBerardinis RJ. Metabolic pathways promoting cancer cell survival and growth. Nat Cell Biol 2015; 17:351–359.
20. Dey P, Li J, Zhang J, Chaurasiya S, Strom A, Wang H, Liao W-T, Cavallaro F, Denz P, Bernard V, Yen E-Y, Genovese G, Gulhati P, Liu J, Chakravarti D, Deng P, Zhang T, Carbone F, Chang Q, Ying H, Shang X, Spring DJ, Ghosh B, Putluri N, Maitra A, Wang YA, DePinho RA. Oncogenic KRAS-Driven Metabolic Reprogramming in Pancreatic Cancer Cells Utilizes Cytokines from the Tumor Microenvironment. Cancer Discov 2020; 10:608–625.

21. Michalski CW, Oti FE, Erkan M, Sauliunaite D, Bergmann F, Pacher P, Batkai S, Müller MW, Giese NA, Friess H, Kleeff J. Cannabinoids in pancreatic cancer: Correlation with survival and pain. Int J Cancer 2008; 122:742–750.
22. Harismendy O, Bansal V, Bhatia G, Nakano M, Scott M, Wang X, Dib C, Turlotte E, Sipe JC, Murray SS, Deleuze JF, Bafna V, Topol EJ, Frazer KA. Population sequencing of two endocannabinoid metabolic genes identifies rare and common regulatory variants associated with extreme obesity and metabolite level. Genome Biol 2010; 11:R118.
23. Xiang W, Shi R, Kang X, Zhang X, Chen P, Zhang L, Hou A, Wang R, Zhao Y, Zhao K, Liu Y, Ma Y, Luo H, Shang S, Zhang J, He F, Yu S, Gan L, Shi C, Li Y, Yang W, Liang H, Miao H. Monoacylglycerol lipase regulates cannabinoid receptor 2-dependent macrophage activation and cancer progression. Nat Commun 2018; 9:2574.
24. Giordano S, Columbano A. Met as a therapeutic target in HCC: Facts and hopes. J Hepatol 2014; 60:442–452.
25. Choueiri TK, Heng DYC, Lee JL, Cancel M, Verheijen RB, Mellemgaard A, Ottesen LH, Frigault MM, L'Hernault A, Szijgyarto Z, Signoretti S, Albiges L. Efficacy of Savolitinib vs Sunitinib in Patients With MET-Driven Papillary Renal Cell Carcinoma. JAMA Oncol 2020; 6:1247–1255.
26. Gijón MA, Riekhof WR, Zarini S, Murphy RC, Voelker DR. Lysophospholipid Acyltransferases and Arachidonate Recycling in Human Neutrophils*. J Biol Chem 2008; 283:30235–30245.
27. Ding F, Gao F, Zhang S, Lv X, Chen Y, Liu Q. A review of the mechanism of DDIT4 serve as a mitochondrial related protein in tumor regulation. Sci Prog 2021; 104:0036850421997273.
28. Almeida R, Amado M, David L, Levery SB, Holmes EH, Merkx G, Kessel AG van, Rygaard E, Hassan H, Bennett E, Clausen H. A Family of Human β4-Galactosyltransferases CLONING AND EXPRESSION OF TWO NOVEL UDP-GALACTOSE:β-N-ACETYLGLUCOSAMINE β1,4-GALACTOSYLTRANSFERASES, β4Gal-T2 AND β4Gal-T3*. J Biol Chem 1997; 272:31979–31991.
29. Ahmad I, Hoessli DC, Walker-Nasir E, Choudhary MI, Rafik SM, Shakoori AR, Nasir-ud-Din. Phosphorylation and glycosylation interplay: Protein modifications at hydroxy amino acids and prediction of signaling functions of the human β3 integrin family. J Cell Biochem 2006; 99:706–718.
30. Scott S, Collet J, Baber U, Yang Y, Peter I, Linderman M, Sload J, Qiao W, Kini A, Sharma S, Desnick R, Fuster V, Hajjar R, Montalescot G, Hulot J. Exome sequencing of extreme clopidogrel response phenotypes identifies B4GALT2 as a determinant of on-treatment platelet reactivity. Clin Pharmacol Ther 2016; 100:287–294.
31. Halbrook CJ, Lyssiotis CA. Employing Metabolism to Improve the Diagnosis and Treatment of Pancreatic Cancer. Cancer Cell 2017; 31:5–19.
32. Busnelli M, Manzini S, Parolini C, Escalante-Alcalde D, Chiesa G. Lipid phosphate phosphatase 3 in vascular pathophysiology. Atherosclerosis 2018; 271:156–165.
33. Choi R, Ham JR, Lee H, Cho HW, Choi M, Park S, Lee J, Kim M, Seo K, Lee M. Scopoletin Supplementation Ameliorates Steatosis and Inflammation in Diabetic Mice. Phytotherapy Res 2017; 31:1795–1804.

34. Xu Y, Jin Y, Gao S, Wang Y, Qu C, Wu Y, Ding N, Dai Y, Jiang L, Liu S. Prognostic Signature and Therapeutic Value Based on Membrane Lipid Biosynthesis-Related Genes in Breast Cancer. J Oncol 2022; 2022:7204415.

Figures



Ontology	ID	Description
KEGG	hsa01200	Carbon metabolism
KEGG	hsa00010	Glycolysis / Gluconeogenesis
KEGG	hsa01230	Biosynthesis of amino acids
KEGG	hsa00520	Amino sugar and nucleotide sugar metabolism
KEGG	hsa00030	Pentose phosphate pathway

Figure 1

(A): The flowchart of the screening of differentially expressed genes (DEGs). **(B):**Ten postoperative tissues of PC patients who were monitored by ultrasonography were performed with RNA bulk sequencing, and then the differentially expressed genes analysis was performed according to the stained intensity of CD8+ T cells in corresponding patients. **(C)**: The postoperative PC tissues of thirteen patients who undergone the Whipple-procedure operation were performed with single cell RNA sequencing, and then the differentially expressed genes analysis was performed according to the stained intensity of CD8+ T cells in corresponding patients. **(D):** the differentially expressed genes analysis was performed according to the proportion of CAFs in corresponding ten patients.**(E)**: The intersection of DEGs was shown with venn diagram; all differentially expressed genes were selected according to p.adj value (p.adj<0.05) and LogFC (LogFC>1.5). **(F)**: KEGG analysis was conducted with the DEGs. **(G)**: GO analysis was conducted with the DEGs.

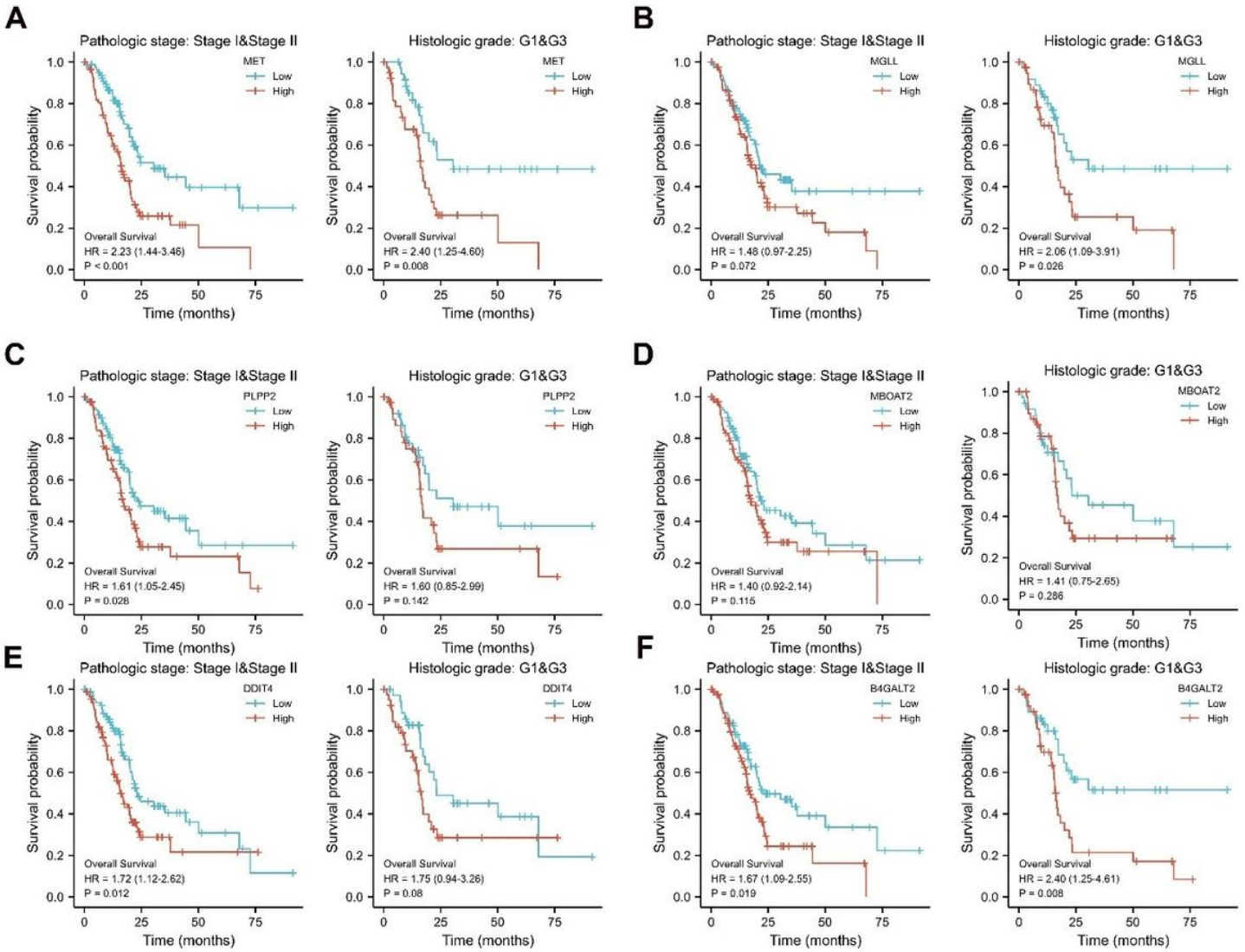


Figure 2

(A-F) the Kaplan–Meier analysis was performed with the metabolism-related gene signature in different clinicopathological characteristics of pancreatic cancer (Pathological stage and Histological grade); A, MET; B, MGLL; C, PLPP2; D, MBOAT2; E, DDIT2; F, B4GALT2.

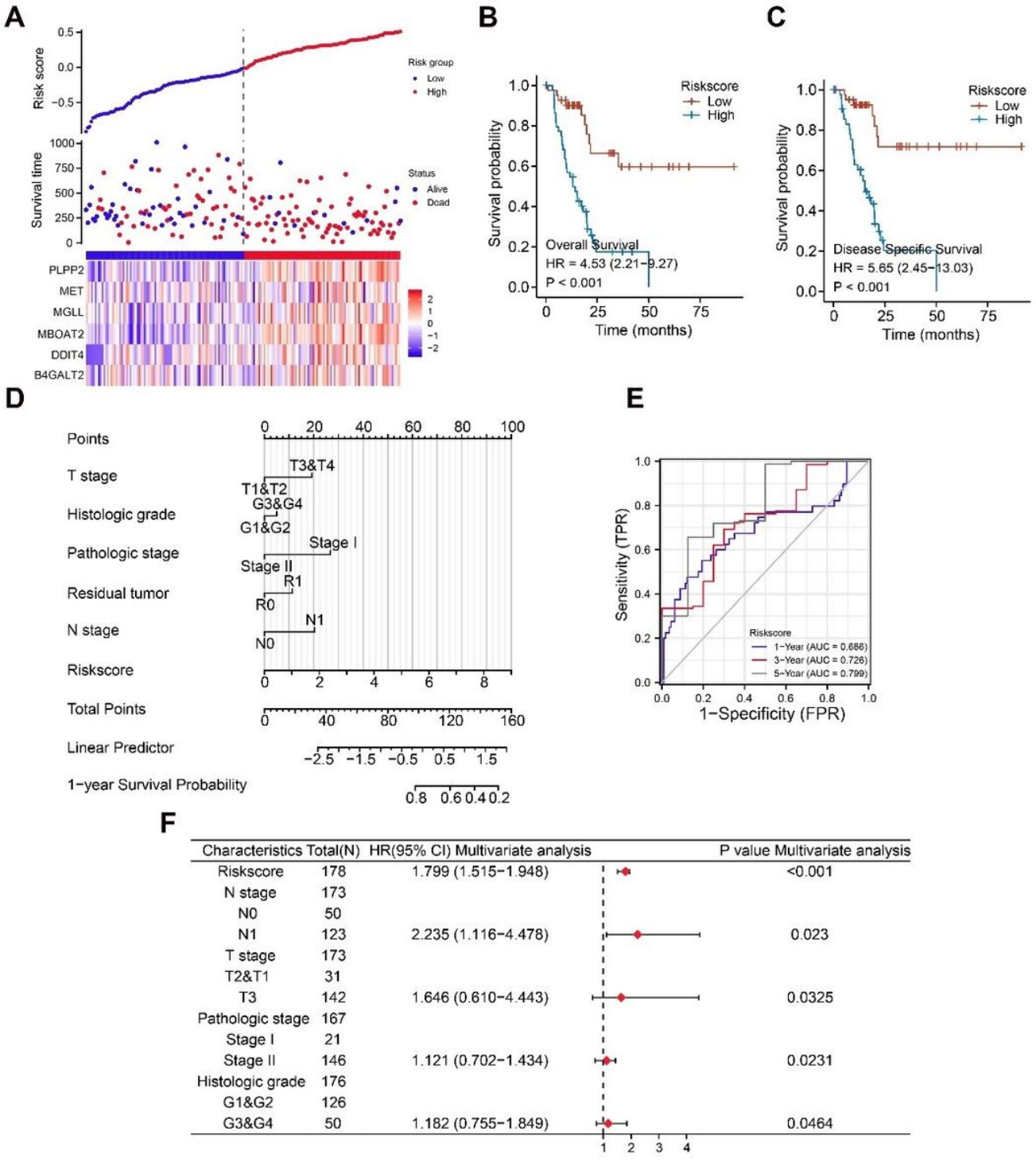


Figure 3

(A): The correlation analysis among these 6 genes was conducted and graphically represented with heatmap; The overall survival rate of different MRS subgroup was calculated according to risk model; the expression level of eight genes in the model was displayed. **(B-C)**: The overall survival (OS) and disease free survival (DFS) analysis were conducted between high risk score group and low risk score group. **(D)**: A receiver operating characteristic (ROC) curve was used for describing a prognostic clinical feature of

MRM. **(E)**: The nomo model was constructed with varieties of factors containing MRM. **(F)**: Multi-variate cox analysis was performed with critical clinical features; hazard ratio (HR)>1; P value<0.05.

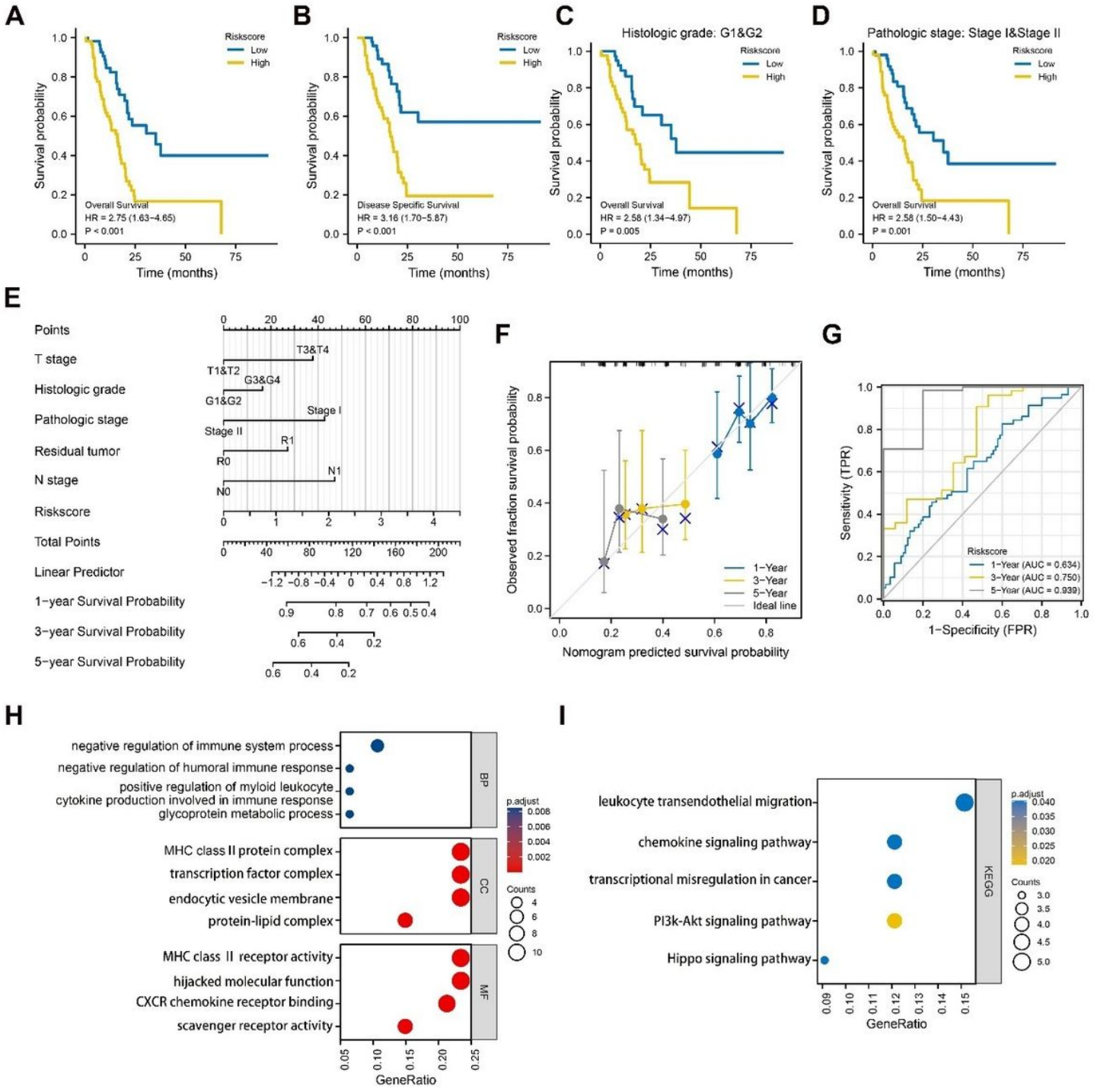


Figure 4

(A-B): The external validation of risk model with partial data from dataset GSE62452; The Kaplan–Meier (KM) survival analysis of disease-free survival (DFS; no recurrence/progression) and overall survival (OS) between high risk score group and low risk score group was performed. **(C)**: K-M survival analysis in

subgroups of clinical characteristics (histological grade, pathologic stage) was conducted between high risk score group and low risk score group. **(D)**: A nomogram consisting of the variables associated with OS was constructed; **(E)**: The calibration curve showed good performance consistent with the nomogram's 1-3- or 5-year OS estimates. **(F)**: A receiver operating characteristic (ROC) curve was used for describing a prognostic clinical feature of risk score. **(G-H)**: the differentially up-regulated expressed genes in the high risk score group were conducted with KEGG and GO analysis; P. adjust <0.05.

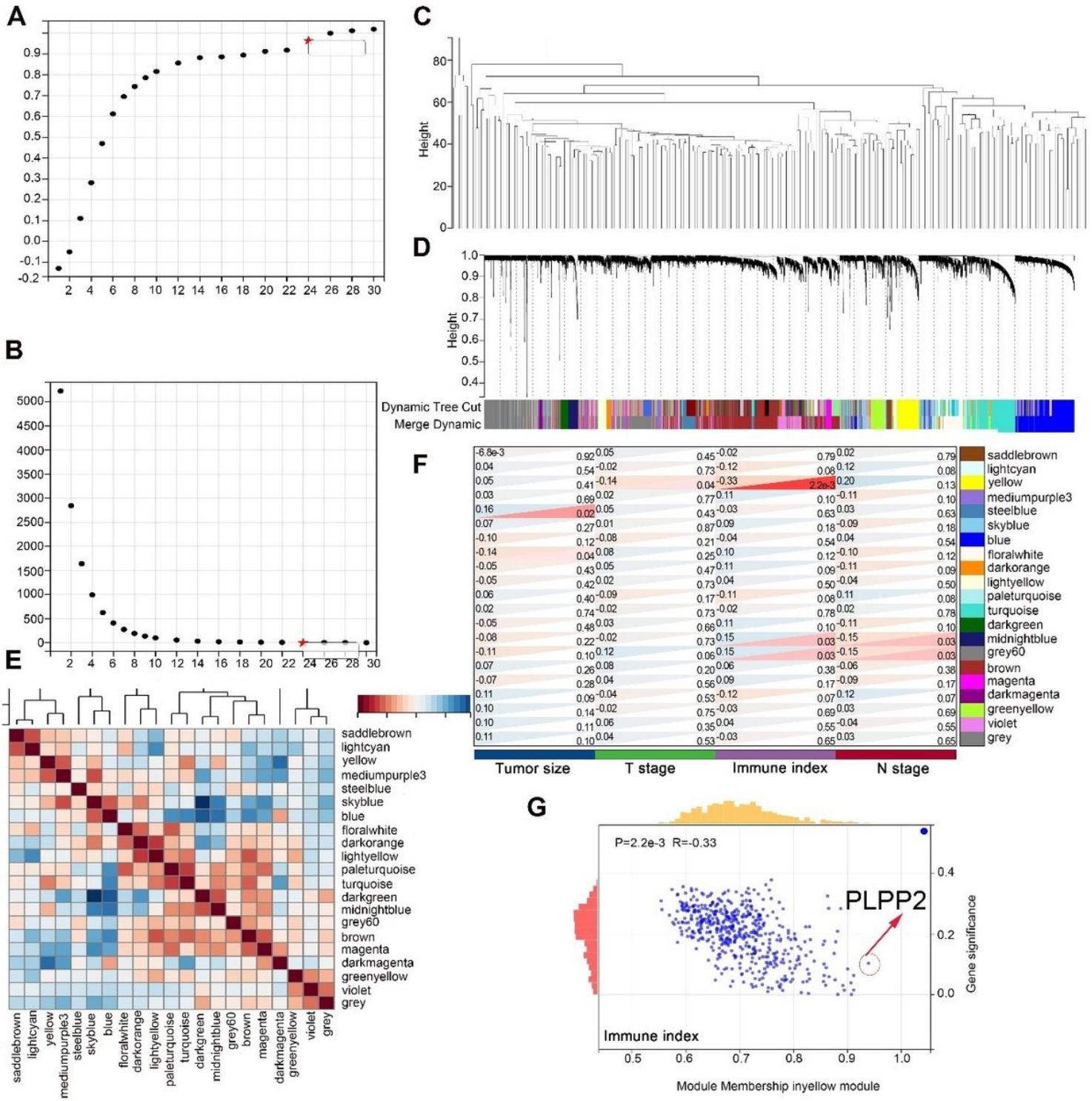


Figure 5

(A-B): The power of β was calculated and selected as the soft threshold to establish the scale-free network. **(C)**: A total of 178 samples with clinical characteristics were included in WGCNA and were clustered by machine learning. **(D)**: Variable co-expressed modules were identified after removing the gray modules by combined dynamic tree cutting. **(E)**: The relationship between variable modules and immune score was calculated respectively. **(F)**: A scatter plot was mapped between GS and MM (colored modules and phenotypes) in PC patients. **(G)**: The targeted gene was considered to be one of the critical genes between the risk model and weighted co-expressed network through Venn diagram.

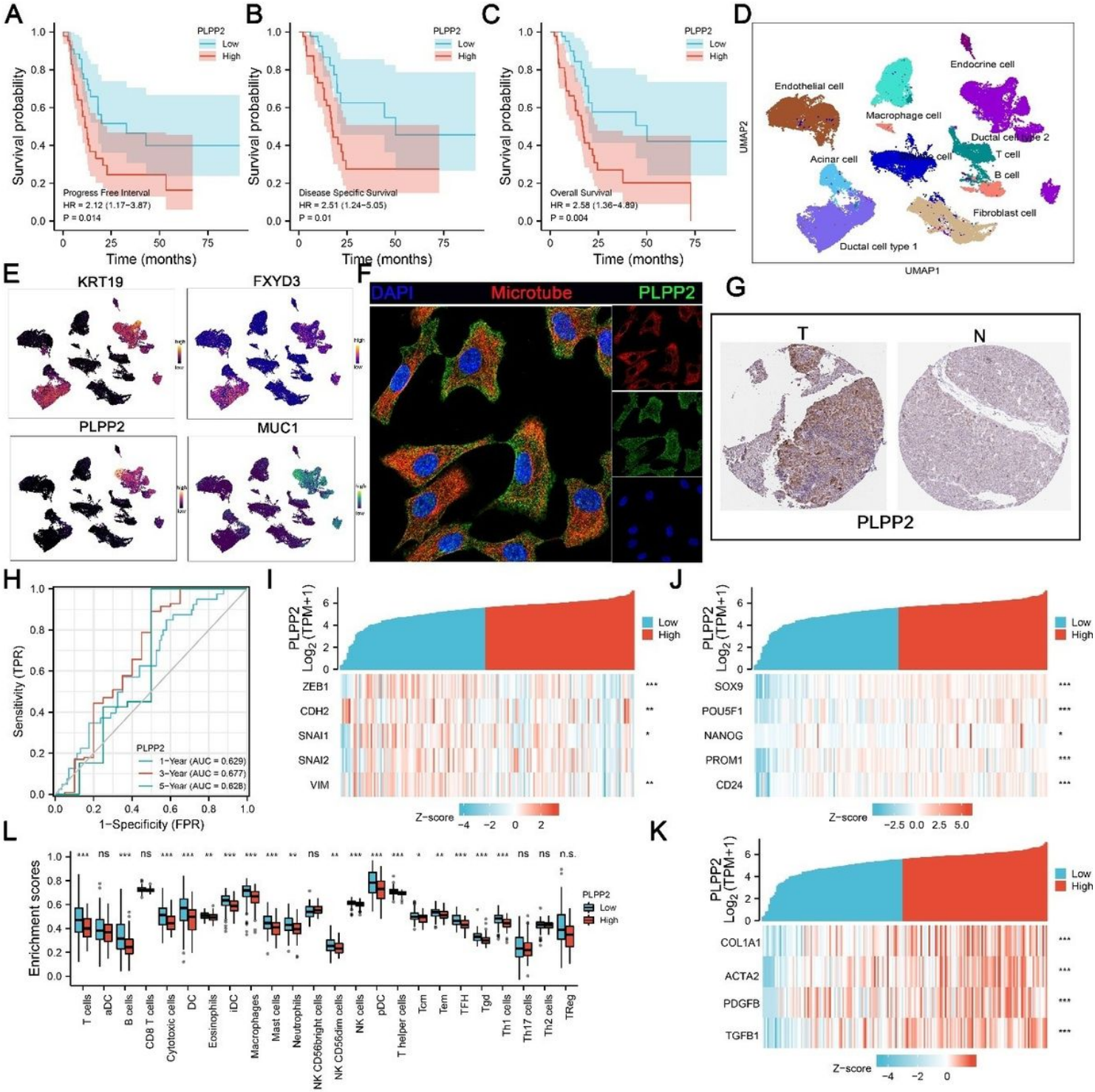


Figure 6

(A-C): In the data set of 178 cases of pancreatic cancer in TCGA database, we performed the survival and prognosis analysis in PC patients with differential expression of PLPP2. **(D-E)**: In order to further explore the function of PLPP2 in pancreatic cancer, we analyzed the characteristics of PLPP2 expression in the ductal epithelial subgroup using the single cell transcriptome dataset in GEO database (GSE12345). KRT19 and EPCAM were used as markers for the normal ductal epithelium; MUC1 and FXYD3 were used to verify the malignant ductal cells. **(F-G)**: We explored the subcellular localization of PLPP2 in pancreatic cancer cells through the *human protein atalas* website, and the expression of PLPP2 in cancer and precancerous tissues from the immunohistochemical level. **(H)**: The predictive efficiency of PLPP2 for the survival rates of pancreatic cancer was performed by timeROC analysis. **(J-K)**: The spearman's correlation analysis was performed between PLPP2 molecules and classical stemness genes (SOX9, POU5F1, NANOG, PROM1). EMT-related genes (ZEB1, SNAI1, SNAI2, VIM) and Activated PSCs-related genes (COL1A1, ACTA2, PDGFB, TGFB1). A significant positive correlation among them was presented with co-expressed heatmap, and the correlation was statistically significant. **(L)**: The correlation analysis of immune infiltration in patients was conducted between high- and low- PLPP2 groups.

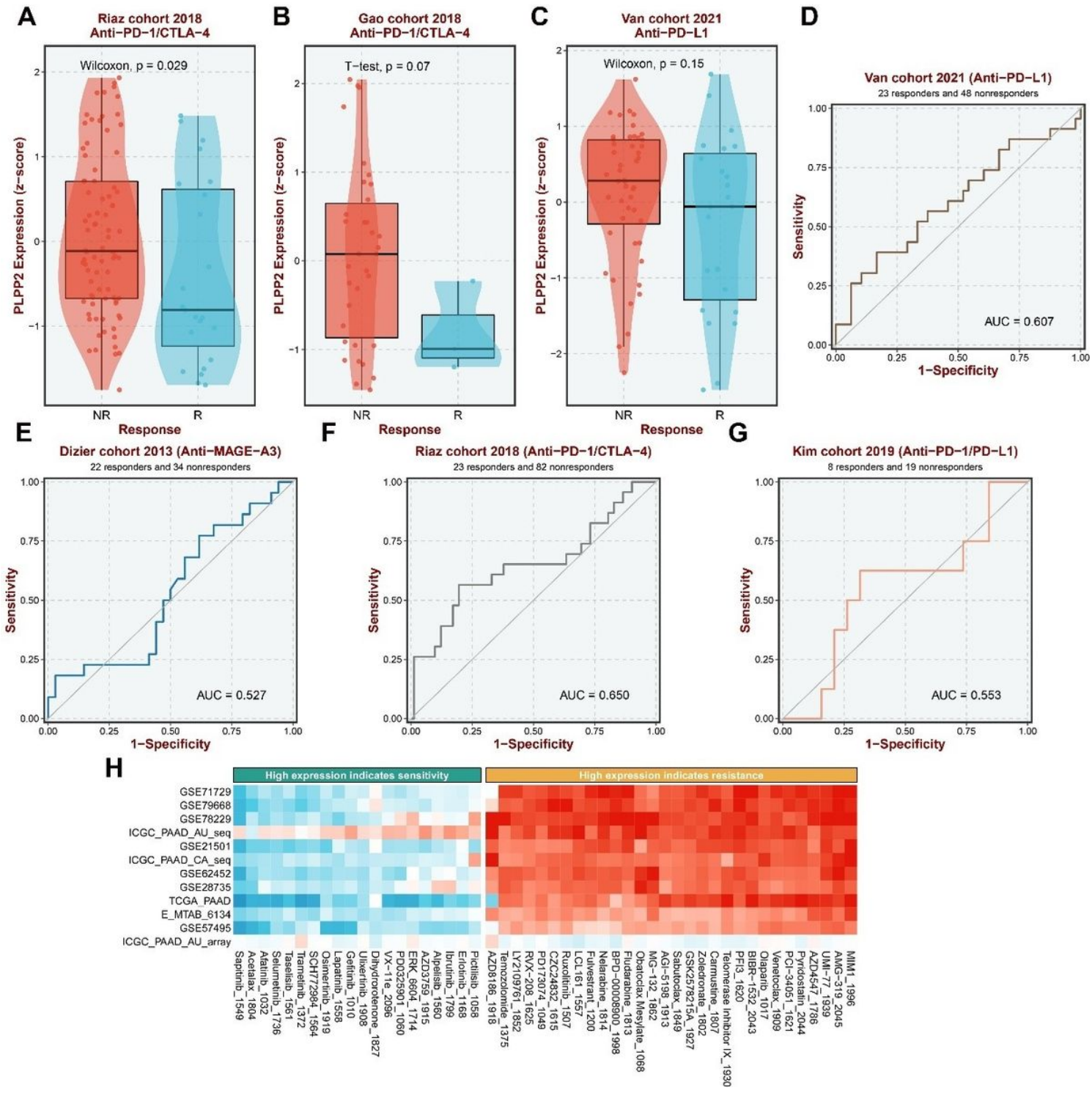


Figure 7

(A-C): The effect of PLPP2 on immune checkpoint inhibitor therapy in Riza cohort 2018(Anti-PD-1/CTLA-4), Gao cohort 2018(Anti-PD-1/CTLA-4), Van cohort 2021(Anti-PD-L1) datasets. **(D-G)**: The receiver operating characteristic curve (ROC) shown that the effect of PLPP2 on immune checkpoint inhibitors in Dizier cohort 2013, Van cohort 2021, Riaz cohort 2018 and Kim cohort 2019. **(H)**: The effect on drug sensitivity analysis in two obvious datasets with different expression of PLPP2.

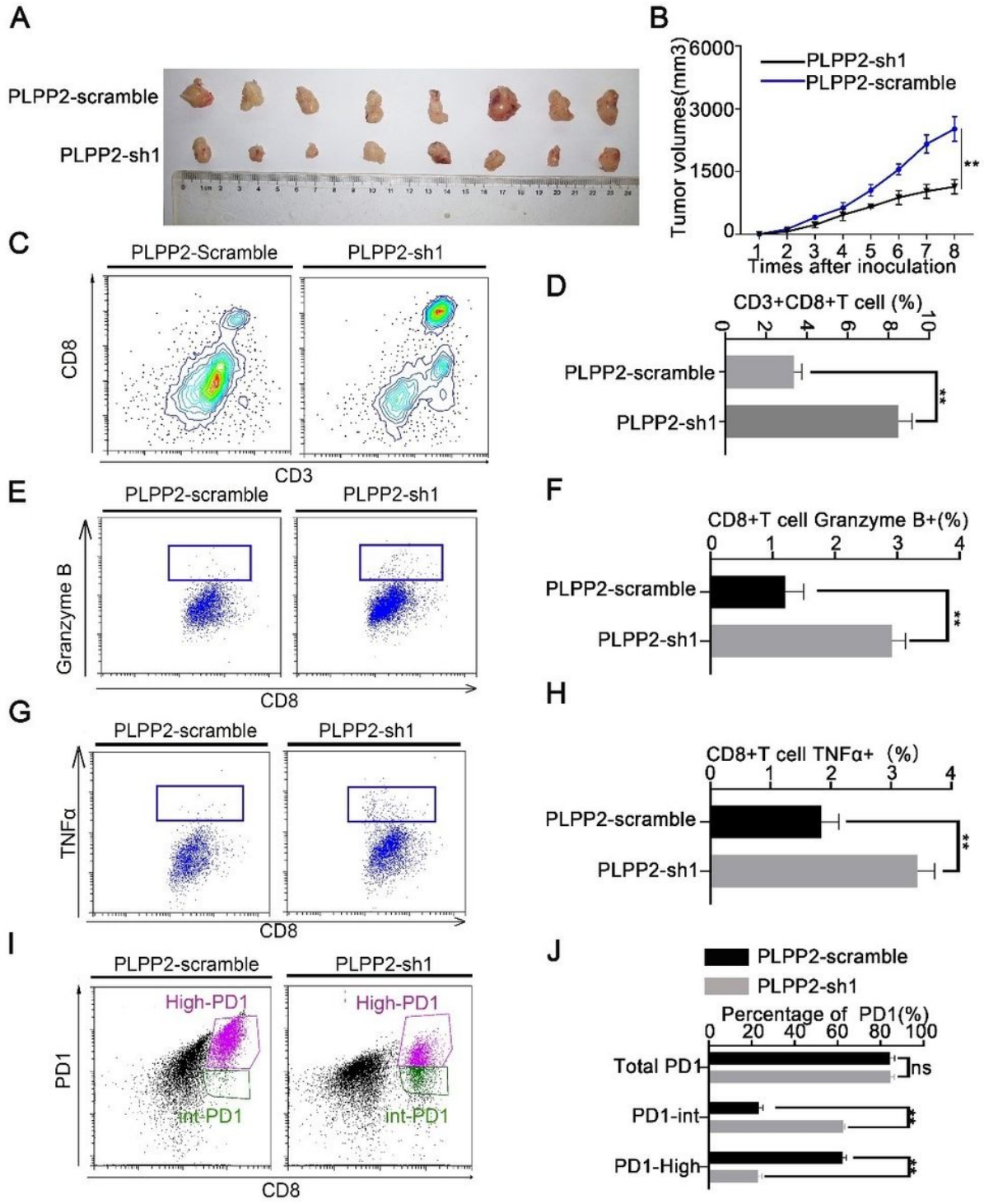


Figure 8

(A-B): KPC-PLPP2-scramble/sh cell lines were implanted subcutaneously in C57BL/6 immunocompetent mice. Tumors implanted subcutaneously were detected in vivo by vernier caliper three times a week; statistical analysis of the fold change of tumor size was calculated. **(C-D)**: Subcutaneous tumors of C57BL/6 immunocompetent mice were harvested after 4 weeks, and tumor tissues were taken for flow cytometry to detect CD8+T cell infiltration ratio, the Y axis represents CD8 (CD45 +) gated, and the X axis represents CD3; statistical analysis was shown by histogram; the mean + SD was used for statistical

analysis between groups, * *, P < 0.01; n.s., no significant statistical difference. **(E-H)**: Flow cytometry was used to detect the infiltration ratio of CD8 + T cells / TNFα + (E-F) and CD8 + T cells / Grzmb + (G-H), the Y axis represents CD8 + T cell, and the X axis represents TNFα / Grzmb , the data were statistically analyzed by histogram; the mean + SD was used for statistical analysis between groups, * *, P < 0.01 ; n. s., no significant statistical difference ; **(I-J)**: The infiltration ratio of CD8/ PD-1 + T cells was detected by flow cytometry ; the Y axis represents CD8 + T cell, and the X axis represents total / int / high /PD-1+; the mean + SD was used for statistical analysis between groups, * *, P < 0.01 ; n. s., no significant statistical difference.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- SupplementaryFigure1.jpg
- SupplementaryFigure2.jpg
- SupplementaryFigure3.jpg
- Supplementarytable1.docx
- Supplementarytable2.docx