



# Glycolysis-Related Gene Expression Profiling Screen for Prognostic Risk Signature of Pancreatic Ductal Adenocarcinoma

Wenjing Song<sup>1</sup>, Xin He<sup>1</sup>, Pengju Gong<sup>1</sup>, Yan Yang<sup>1</sup>, Sirui Huang<sup>1</sup>, Yifan Zeng<sup>1</sup>, Lei Wei<sup>2</sup> and Jingwei Zhang<sup>1\*</sup>

<sup>1</sup> Department of Breast and Thyroid Surgery, Hubei Key Laboratory of Tumor Biological Behaviors, Hubei Cancer Clinical Study Center, Zhongnan Hospital, Wuhan University, Wuhan, China, <sup>2</sup> Department of Pathology and Pathophysiology, School of Basic Medical Sciences, Wuhan University, Wuhan, China

## OPEN ACCESS

### Edited by:

Alessandro Romanel,  
University of Trento, Italy

### Reviewed by:

Paolo Gandellini,  
University of Milan, Italy  
Xiaojie Zhao,  
Rosalind Franklin University of  
Medicine and Science, United States

### \*Correspondence:

Jingwei Zhang  
zjwzhang68@whu.edu.cn

### Specialty section:

This article was submitted to  
Cancer Genetics and Oncogenomics,  
a section of the journal  
Frontiers in Genetics

Received: 14 February 2021

Accepted: 25 May 2021

Published: 23 June 2021

### Citation:

Song W, He X, Gong P, Yang Y,  
Huang S, Zeng Y, Wei L and Zhang J  
(2021) Glycolysis-Related Gene  
Expression Profiling Screen for  
Prognostic Risk Signature of  
Pancreatic Ductal Adenocarcinoma.  
Front. Genet. 12:639246.  
doi: 10.3389/fgene.2021.639246

**Objective:** Pancreatic ductal adenocarcinoma (PDAC) is highly lethal. Although progress has been made in the treatment of PDAC, its prognosis remains unsatisfactory. This study aimed to develop novel prognostic genes related to glycolysis in PDAC and to apply these genes to new risk stratification.

**Methods:** In this study, based on the Cancer Genome Atlas (TCGA) PAAD cohort, the expression level of glycolysis-related gene at mRNA level in PAAD and its relationship with prognosis were analyzed. Non-negative matrix decomposition (NMF) clustering was used to cluster PDAC patients according to glycolytic genes. Prognostic glycolytic genes, screened by univariate Cox analysis and LASSO regression analysis were established to calculate risk scores. The differentially expressed genes (DEGs) in the high-risk group and the low-risk group were analyzed, and the signal pathway was further enriched to analyze the correlation between glycolysis genes. In addition, based on RNA-seq data, CIBERSORT was used to evaluate the infiltration degree of immune cells in PDAC samples, and ESTIMATE was used to calculate the immune score of the samples.

**Results:** A total of 319 glycolysis-related genes were retrieved, and all PDAC samples were divided into two clusters by NMF cluster analysis. Survival analysis showed that PDAC patients in cluster 1 had shorter survival time and worse prognosis compared with cluster 2 samples ( $P < 0.001$ ). A risk prediction model based on 11 glycolysis genes was constructed, according to which patients were divided into two groups, with significantly poorer prognosis in high-risk group than in low-risk group ( $P < 0.001$ ). Both internal validation and external dataset validation demonstrate good predictive ability of the model ( $AUC = 0.805$ ,  $P < 0.001$ ;  $AUC = 0.763$ ,  $P < 0.001$ ). Gene aggregation analysis showed that DEGs highly expressed in high-risk group were mainly concentrated in the glycolysis level, immune status, and tumor cell proliferation, etc. In addition, the samples in high-risk group showed immunosuppressed status and infiltrated by relatively more macrophages and less CD8+T cell.

**Conclusions:** These findings suggested that the gene signature based on glycolysis-related genes had potential diagnostic, therapeutic, and prognostic value for PDAC.

**Keywords:** glycolysis, gene signature, overall survival, prognostic biomarker, pancreatic ductal adenocarcinoma

## INTRODUCTION

Pancreatic ductal adenocarcinoma (PDAC) is the most common type of pancreatic cancer, accounting for about 90% of pancreatic cancer cases. PDAC is a highly fatal cancer with a 5-year overall survival rate of 6% (2–9%) (Ilic and Ilic, 2016). By 2030, PDAC is projected to be the second leading cause of cancer-related deaths (Rahib et al., 2014). Surgical resection is considered to be the only treatment for PDAC, which can significantly prolong the survival time. However, due to the hidden onset, rapid progression, early metastasis and lack of sensitive screening methods (Singhi et al., 2019), most PDAC are confirmed in its advanced stage and lost the thus become unresectable. Therefore, about 80% of people diagnosed with PDAC will die within a year and thus early diagnosis of PDAC is particularly important (Zhang et al., 2017). For diagnostic methods, the specificity and sensitivity of tumor biomarkers for predicting PDAC are not satisfactory (Zhou et al., 2015), and it is difficult to find small lesions by computed tomography (CT) (Fox et al., 2016; Zhang et al., 2017). Although pathological examination is the gold standard for the diagnosis of PDAC, the deep location of pancreas limits its application. Therefore, it is urgent to develop a method and strategy for early detection of PDAC, which will be beneficial to the diagnosis and treatment of PDAC patients. In addition, radiotherapy, chemotherapy, targeted molecular therapy and immunotherapy have also been shown to be effective in the treatment of PDAC (Kamisawa et al., 2016). The diversity of treatment options requires more individualized management of PDAC patients, which can improve both the treatment of cancer and the quality of life. Tumor progression and drug response in PDAC patients are closely related to the molecular characteristics, phenotypic differences and tumor microenvironment (TME). According to these characteristics, different classification systems for PDAC subtypes could be established to predict the prognosis of patients and select therapeutic drugs and therapies (Collisson et al., 2011; Moffitt et al., 2015; Bailey et al., 2016; Puleo et al., 2018). Advances in tumor molecular biology have led to the development of new predictive tools based on prognostic genes. These prognostic markers reflecting tumor progression at the molecular level may contribute to more accurate personalized survival prediction.

Microenvironment is the cellular environment where cells grow, proliferate and invade, and tumor cells are no exception (Wang et al., 2018). Most solid tumors rely heavily on aerobic glycolysis for energy production due to the metabolic reprogramming of tumor cells to facilitate the aerobic glycolysis process to adapt to their heterogeneous microenvironment. In addition to the tumor cells, the TME also includes the surrounding immune cells, fibroblasts and immune cells. Because of the dense connective tissue and vascular microenvironment of PDAC, PDAC cells are difficult to penetrate and in a low-perfusion environment, which promotes metabolic rearrangement in the PDAC so that tumor cells can make full use of oxygen even in a hypoxia state (Fu et al., 2018; Weniger et al., 2018). Thus, PDAC generally displays enhanced glycolysis, including overexpression of glycolytic enzymes and increased lactic acid production, which is caused by mitochondrial

dysfunction, abnormal expression of oncogenic genes, specific transcription factors, hypoxic tumor microenvironment, and tumor-associated macrophage (Cheng et al., 2019; Karasinska et al., 2020; Yang et al., 2020). This energy metabolic pathway not only provides energy to cancer cells, but also produces metabolic intermediates that promote cell proliferation, invasion and drug resistance of cancer (Deberardinis et al., 2008). Reprogramming of metabolism in cancer cells is regulated by multiple factors and signaling pathways, such as hypoxia inducible factor (HIF-1), Myc, p53, and the PI3K/Akt/mTOR pathway (Pelicano et al., 2006; Dang et al., 2009; Liao et al., 2009; Masui et al., 2013). Some preclinical and clinical studies have shown that drugs targeting these factors and signaling pathways and the use of anti-glycolysis agents to deprive the basic metabolic needs of cancer cells and interfere with cancer growth are effective as therapies to inhibit cancer progression (Abdel-Wahab et al., 2019; Jagust et al., 2019). In addition, it is gratifying to note that anti-glycolysis agents have been found to have the potential to increase the sensitivity of cancer cells and improve treatment resistance. Therefore, it is of great significance to search for molecules related to glycolysis and explore their expression characteristics and functional involvement in PDAC.

In the present study, 11 prognostic glycolysis-related genes were selected based on high-throughput sequencing results and clinicopathological features of the PDAC data sets, and the prognostic gene signature was proposed, which was used to describe the glycolytic level in PDAC samples. Multivariate Cox regression confirmed that prognostic gene signature was independent prognostic influential factor of PDAC, and was validated using GEO data sets. In addition, we described the gene expression at the protein level, and investigated in depth the correlation between the signal pathways involved in the gene signature and biochemical processes and tumor immunity. These results may be helpful for the study of glycolysis-related molecules in PDAC.

## METHODS

### Patient Data Acquisition

The Cancer genome atlas (TCGA) (<https://portal.gdc.cancer.gov/>) is a cancer research project established by the national cancer institute and the national human genome research institute jointly, covering 33 kinds of cancer. Download RNA-seq and corresponding clinical data of PDAC patients (survival time, survival status, diagnostic age, gender, history of smoking, history of alcohol, history of diabetes, history of chronic pancreatitis, tumor site, histological grade, pathological T, N, M, stage, residual tumor and radiation therapy). After data cleaning, a total of 173 patients with PDAC had complete survival data, and 73 patients had complete clinicopathological data. Data download and online analysis has been commenced on 1 November 2020 (TCGA: V21.0).

### Identification of Glycolysis-Related Genes

The Molecular Signature Database (MSigDB) (<https://www.gsea-msigdb.org/gsea/msigdb/index.jsp>) (Subramanian et al., 2005; Liberzon et al., 2011, 2015) provides a collection of annotated

gene sets to analyze. Search for “glycolysis” by keyword, download all glycolysis-related data sets, extract and sort out the contained genes. Then, we performed Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analyses on these 319 glycolysis-related genes to verify whether these genes are involved in regulating the process of glycolysis.

## PDAC Subclasses Identification Based on Glycolysis-Related Genes

The obtained 319 glycolysis-related genes were used for non-negative matrix factorization (NMF) clustering. The purpose of NMF was to identify potential features in the gene expression profile by decomposing the original matrix into different non-negative matrix. Use the “NMF” R package to perform unsupervised NMF clustering with 1,000 repeat samples and a maximum grouping of 6 on the metadata set. The cumulative distribution function (CDF) and consensus heatmap were used to evaluate the optimal  $k$  value, and the TCGA-PAAD samples were divided into different clusters according to the glycolysis level of tumor tissues.

## Prognosis Analysis

Kaplan-Meier (K-M) method was used to plot the survival curves between the expression level of 319 glycolysis-related genes and the overall survival time (OS) of PDAC patients (cutoff by the median expression level). The significance of the difference was tested by log-rank test, and the genes whose  $P < 0.01$  were screened for subsequent study. Differences in OS, histological grade and pathological stage between the two groups were compared and visualized by box plot using R language (version 3.6.1).

## Gene Signature Identification and Score Construction

The prognostic related genes were identified by univariate Cox regression analysis. After that, least absolute shrinkage and selection operator (LASSO) regression was employed to identify independent prognostic influencing genes powerfully associated ( $P < 0.05$ ) with OS in PDAC patients. The risk score was calculated by the following formula:

$$\text{Risk score} = \sum_{i=1}^n \text{Coef}(i) X(i)$$

What needs to be commented is that “ $n$ ” represents the count of genes in the model, “ $\text{Coef}(i)$ ,” that is “coefficient,” represents the coefficient of each gene,  $X(i)$  means the mRNA expression level of each gene. When the coefficient is  $>0$ , it indicates that the overexpression of this gene increases the risk of PDAC patients, otherwise, it means that the gene has a protective effect on patients with PDAC. Download the Gene Expression Omnibus (GEO) data set (GSE62452) (Yang et al., 2016) to externally validate the model. Patients in TCGA-PAAD and GEO were divided into low-risk group and high-risk group according to the median risk level, respectively. The survival differences of the two groups were compared and visualized with the survival status plot, risk heatmap and survival curve. The area under the curve (AUC) of the 95% confidence interval was determined

according to the receiver operating characteristic curve (ROC), and the accuracy and specificity of the model were quantitatively evaluated. Diagnosis of age, gender, history of smoking, history of alcohol, history of diabetes, history of chronic pancreatitis, tumor site, histological grade, pathological T, N, M, stage, residual tumor, radiation therapy and risk score were all included in this study for univariate and multivariate Cox regression analysis, which determined that risk score based on 11 glycolysis-related genes was the independent prognostic factor for PDAC.

## Functional Inference

Gene Set Enrichment Analysis (GSEA), a desktop software used to analyze gene sets, can be downloaded from the Broad Institute GSEA website (<https://www.gsea-msigdb.org/gsea/index.jsp>) (Subramanian et al., 2005). To infer functional annotations of 11 glycolysis-related genes, GO enrichment [c5.all.v7.2.symbols.gmt (Gene oncology)] and KEGG pathway analysis [c2.cp.kegg.v7.2.symbols.gmt (Curated)] of differentially expressed genes (DEGs) in low-risk group and high-risk group were enriched by GSEA (version 4.0.1). CHIP - X Enrichment Analysis Version 3 (ChEA3) (<https://maayanlab.cloud/chea3/#top>) (Keenan et al., 2019) is a web-based enrichment tool of transcription factor (TF) analysis. We predicted the top 25 TF that are most closely related to the 11 glycolysis-related genes by using Fisher’s precise test.

## Analysis of Tumor Infiltrate Immune Cells in PDAC

Based on RNA-seq data, the “Cibersort” R package was used to estimate the abundance of 22 TIICs in a single sample (Newman et al., 2019), and the stromal, immune and estimated scores were calculated using the “ESTIMATE” R package (Yoshihara et al., 2013). The infiltration level and immune scores were compared between the high-risk group and the low-risk group.

## Immunohistochemical Analysis of Glycolysis-Related Genes in PDAC

Human protein mapping (HPA) (<https://www.proteinatlas.org/>) (Uhlén et al., 2015; Uhlen et al., 2017; Thul et al., 2017) of 26,000 human proteins tissue and cell distribution information. In this database, the researchers used highly specific antibodies and immunoassay techniques to examine the expression of proteins in cell lines, tumor tissues and normal samples in detail. The protein expression of 11 genes in PDAC tissues was analyzed using HPA database.

## Statistical Analysis

Most statistical analysis is done through online bioinformatics databases and tools. Other statistical analysis was performed based on R software v3.6.1. When the data is normally distributed, the mean and median of continuous variables are compared by Student’s  $t$ -test, otherwise, use Wilcoxon inspection. The Chi-square test and Fisher test was used to compare clinical and pathological parameters and other categorical variables. Survival rates were assessed using Kaplan-Meier curves and the log-rank test, and univariate and multivariate Cox regression were used to analyze the independent parameters associated with

the OS. All tests were bilateral, and  $P < 0.05$  was considered statistically significant. Pearson coefficient of correlation was calculated to measure the correlation between two variables.

## RESULTS

### Identification of Glycolysis-Related Genes in TCGA-PAAD

A total of 13 glycolysis-related genes sets were retrieved (Table 1). These 13 gene sets contain genes involved in glycolysis, gluconeogenesis and the tricarboxylic acid cycle, metabolism of fructose 1, 6-bisphosphate and fructose 2, 6-bisphosphate, transmembrane transport of lactic acid and other biochemical reactions, which are considered to be up-regulated genes for glycolysis. After summarizing the genes in the above gene sets, 319 glycolysis-related genes were included through deduplication. To verify whether these 319 genes were involved in glycolysis, we conducted in-depth studies using GO and KEGG pathway enrichment analysis. The results showed that the genes enriched in molecular function (MF) term were related to monosaccharide binding and oxidoreductase activity, and that in biological process (BP) term were related to glycolytic metabolism, and the KEGG pathway enrichment analysis involved glycolysis/gluconeogenesis, suggesting that these genes were indeed related to glycolysis (Figures 1A,B).

### PDAC Subclasses Identification Based on Glycolysis-Related Genes

These 319 glycolysis-related genes were used for NMF cluster analysis, and the comprehensive correlation coefficient was used to determine the optimal  $k$  value as 2. TCGA-PAAD samples were then divided into two different clusters, namely cluster 1 ( $n = 101$ ) and cluster 2 ( $n = 72$ ). When  $k = 2$ , the consensus matrix heatmap had clear boundary and minimal interference between subgroups, indicating that the samples had stable clusters (Figures 2A–D). In addition, we compared the differences of survival between two clusters with different glycolytic status. The survival curve showed that patients in cluster 1 had the worse prognosis than those in cluster 2 ( $P = 0.004$ ) (Figure 2E).

In order to better understand the glycolysis status between the samples of the two clusters, the volcano plot of DEGs of the two clusters were shown in Figure 2F, in which 343 genes were up-regulated and 37 genes were down-regulated significantly in the cluster 1 ( $P < 0.05$ ). And the results of gene function and pathway enrichment analysis showed that these DEGs were significantly related to glycolysis and cell proliferation and migration, such as glycolytic process, oxidation-reduction coenzyme metabolic process, oxidoreductase complex, glucose binding, glycolysis, citrate cycle (TCA cycle), cell cycle, autophagy and p53 signaling pathway (Figures 2G,H). In addition, we compared the infiltrating abundance of TIICs between PDAC samples in cluster 1 and cluster 2. The results showed that there were more infiltration of macrophages M0 and less infiltration of B and T cells in

cluster 1 ( $P < 0.01$ ) (Supplementary Figure 1A). Besides, our analysis also found that the samples in cluster 1 had lower immune score, stromal score, and ESTIMATE score ( $P < 0.05$ ), while the tumor purity was higher ( $P < 0.001$ ) (Supplementary Figures 1B–E).

### Gene Signature Identification and Score Construction

The screening process for prognostic glycolysis-related genes was shown in Supplementary Figure 2. The K-M survival curves showed that among the 319 glycolysis-related genes, the expression level of 38 genes was significantly correlated with the OS of PDAC ( $P < 0.01$ ). Among them, 30 genes were significantly up-regulated in PDAC patients with poor prognosis (Supplementary Figure 3), and the high expression of 8 genes may indicate a better prognosis in patients with PDAC (Supplementary Figure 4). In TCGA-PAAD patients, 35 genes associated with total survival were identified by univariate Cox regression analysis ( $P < 0.05$ ) (Table 2). To determine the most powerful prognostic markers, the LASSO regression analysis was used to screen 11 genes and construct risk score signature (Figure 3A) to minimize the risk of overfitting. The risk score of PAAD patients was calculated according to the expression level of gene and regression coefficient, and the results were as follows: Risk score =  $0.00559311^*$  (the expression level of ALDH3B1) +  $0.021443237^*$  (the expression level of PGM1) +  $0.0246051^*$  (the expression level of MET) +  $0.049373852^*$  (the expression level of KIF20A) +  $0.036135631^*$  (the expression level of ABCB6) +  $0.012568672^*$  (the expression level of NT5E) +  $0.512999189^*$  (the expression level of CHST12) +  $0.001510137^*$  (the expression level of GPR87) +  $0.036407687^*$  (the expression level of CDK1) +  $0.000281881^*$  (the expression level of B3GNT3) +  $0.01548257^*$  (the expression level of CACNA1H). LASSO Cox regression fitted 11 most powerful genes into a formula related to prognosis. The PDAC patients were divided into a low-risk group ( $n = 87$ ) and a high-risk group ( $n = 86$ ) based on the median risk score. The expression levels of 11 prognostic markers in the high-risk group and the low-risk group were shown in a box plot (Figure 3B). As can be seen, compared with the low-risk group, ALDH3B1, PGM1, MET, KIF20A, NT5E, GPR87, CDK1, B3GNT3 were significantly up-regulated in the high-risk group ( $P < 0.001$ ), and CHST12, CACNA1H were significantly down-regulated in the high-risk group ( $P < 0.001$ ), while the expression level of ABAC6 was not significantly different between the two groups. Compared with the low-risk group, the high-risk group had more deaths by the time of follow-up, and the distribution of survival status and risk score was shown in Figures 3C,D. The ROC curve showed a good predictive ability of the model ( $AUC = 0.805$ ,  $P < 0.001$ ) based on the gene signature to predict the prognosis of PDAC patients (Figure 3E).

To the diagnosis of age, gender, history of smoking, history of alcohol, history of diabetes, history of chronic pancreatitis, tumor site, histological grade, pathological stage, residual tumor and radiation therapy together with risk score into the univariate and multivariate Cox regression analysis, the results showed that the risk score based on 11 glycolysis-related genes was an

**TABLE 1** | Thirteen glycolysis-related gene sets screened from MSigDB. Standard name: The name of genes sets in MSigDB.

Standard name	Brief description
BIOCARTA FEEDER PATHWAY	Feeder pathways for glycolysis
BIOCARTA GLYCOLYSIS PATHWAY	Glycolysis pathway
HALLMARK GLYCOLYSIS	Genes encoding proteins involved in glycolysis and gluconeogenesis
KEGG GLYCOLYSIS GLUCONEOGENESIS	Glycolysis/gluconeogenesis
MODULE 306	Glycolysis and TCA cycle
REACTOME GLYCOLYSIS	Glycolysis
REACTOME REGULATION OF GLYCOLYSIS BY FRUCTOSE 2,6-BISPHOSPHATE METABOLISM	Regulation of glycolysis by fructose 2,6-bisphosphate metabolism
WP COMPUTATIONAL MODEL OF AEROBIC GLYCOLYSIS	Computational model of aerobic glycolysis
WP GLYCOLYSIS AND GLUCONEOGENESIS	Glycolysis and gluconeogenesis
WP HIF1A AND PPARG REGULATION OF GLYCOLYSIS	HIF1A and PPARG regulation of glycolysis
GO LACTATE TRANSMEMBRANE TRANSPORT	The process in which lactate is transported across a membrane
GO LACTATE TRANSMEMBRANE TRANSPORTER ACTIVITY	Enables the transfer of lactate from one side of a membrane to the other
GO FRUCTOSE 1,6-BISPHOSPHATE METABOLIC PROCESS	The chemical reactions and pathways involving fructose 1,6-bisphosphate

independent prognostic factor in PDAC patients ( $P < 0.001$ , HR = 2.916, 95% CI: 1.627–5.227) (Figures 3E,G).

### Comparison of Clinicopathological Characteristic Among Risk Groups

To further evaluate the impact of risk score on prognosis of PDAC patient, K-M analysis showed that the prognosis of high-risk group was significantly worse than that of low-risk group ( $P < 0.001$ ) (Figure 4A). In addition, we found that the high-risk group included more patients in cluster 1 (cluster 1: cluster 2 = 72:14), and the low-risk group mainly included patients in cluster 2 (cluster 1: cluster 2 = 29:58), the difference was statistically significant ( $P < 0.001$ ) (Figure 4B), indicating that patients in the high-risk group had a more active glycolysis status. Besides, we compared the differences in the pathological stage and histological grade of the cancers between the high-risk group and the low-risk group in detail. We found that although there were no statistically significant differences in T, N, M, Stage and Grade composition between the high-risk group and the low-risk group ( $P > 0.05$ ), the high-risk group tended to include more PDAC samples with late stage and high grade (Figures 4C–G).

### External Validation of the Gene Signature

To further verify the usefulness of the risk score model based on 11 glycolysis-related gene signatures, we downloaded the GSE62452 ( $n = 130$ ) data set from GEO, in which 69 samples were PDAC samples. After removing samples whose gene expression was 0 and samples without survival data, a total of 64 samples were included in the study. According to the above formula, the risk score of the GEO dataset sample was calculated and the patients were divided into the low-risk group ( $n = 32$ ) and the high-risk group ( $n = 32$ ). K-M analysis showed that the high-risk group had a significantly worse prognosis than the low-risk group (Figure 5A). Compared with the low-risk group, the high-risk group had more deaths

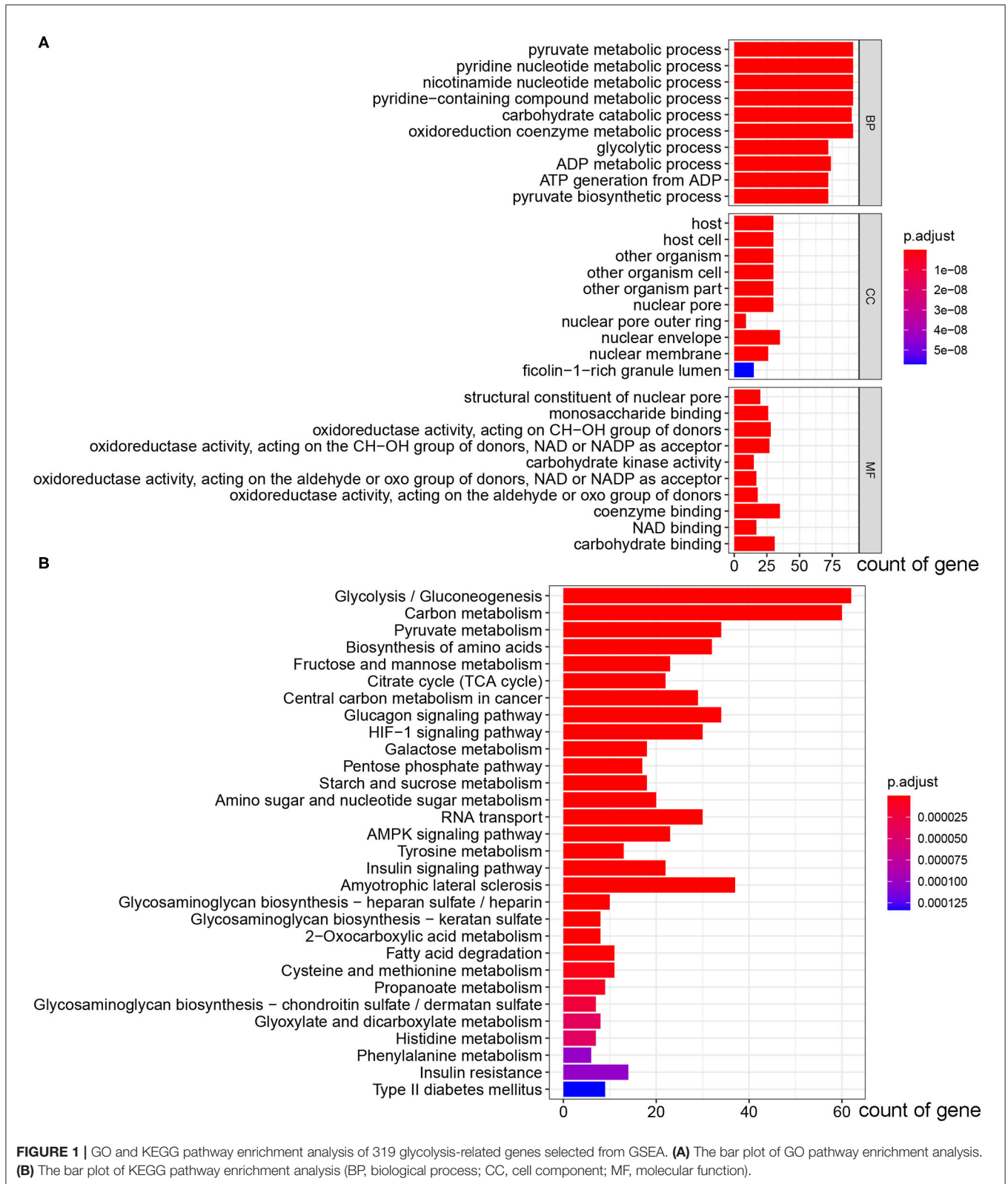
by the time of follow-up. The distribution of survival status and risk score of the patients was shown in Figures 5C,D. Predicting the prognosis of PDAC patients based on this gene signature, the ROC curve showed a good predictive ability of the model (AUC = 0.763,  $P < 0.001$ ) (Figure 5B). The expression levels of 11 prognostic markers in the high-risk group and the low-risk group were shown by heatmap (Figure 5E).

### Function and Mechanism Inference

The DEGs between the high-risk group and the low-risk group were shown in the heatmap (Figure 6A), in which 440 genes were significantly up-regulated and 25 genes were significantly down-regulated in the high-risk group ( $P < 0.05$ ). It is worth noting that GSEA analysis results showed the DEGs were enriched in glycolysis gluconeogenesis, primary immunodeficiency and cell cycle pathway (Figure 6B), which meant that there were differences in glycolysis levels, immune status, and tumor cell proliferation in the high and low risk groups. The correlation plot showed that there were strong positive correlations between MET and other oncogenic glycolysis genes, such as KIF20A and MET ( $r = 0.52961$ ), NT5E and MET ( $r = 0.634197$ ), GPR87 and MET ( $r = 0.522604$ ), etc. ( $P < 0.001$ ), while there were negative correlations between MET and ABCB6, CHST12 and CACNA1H (Figure 6C). Figure 6D depicted a network of top-25 TF with strong regulatory relationships among 11 genes, including HIF-1A.

### Protein Expression Analysis of Glycolysis-Related Genes

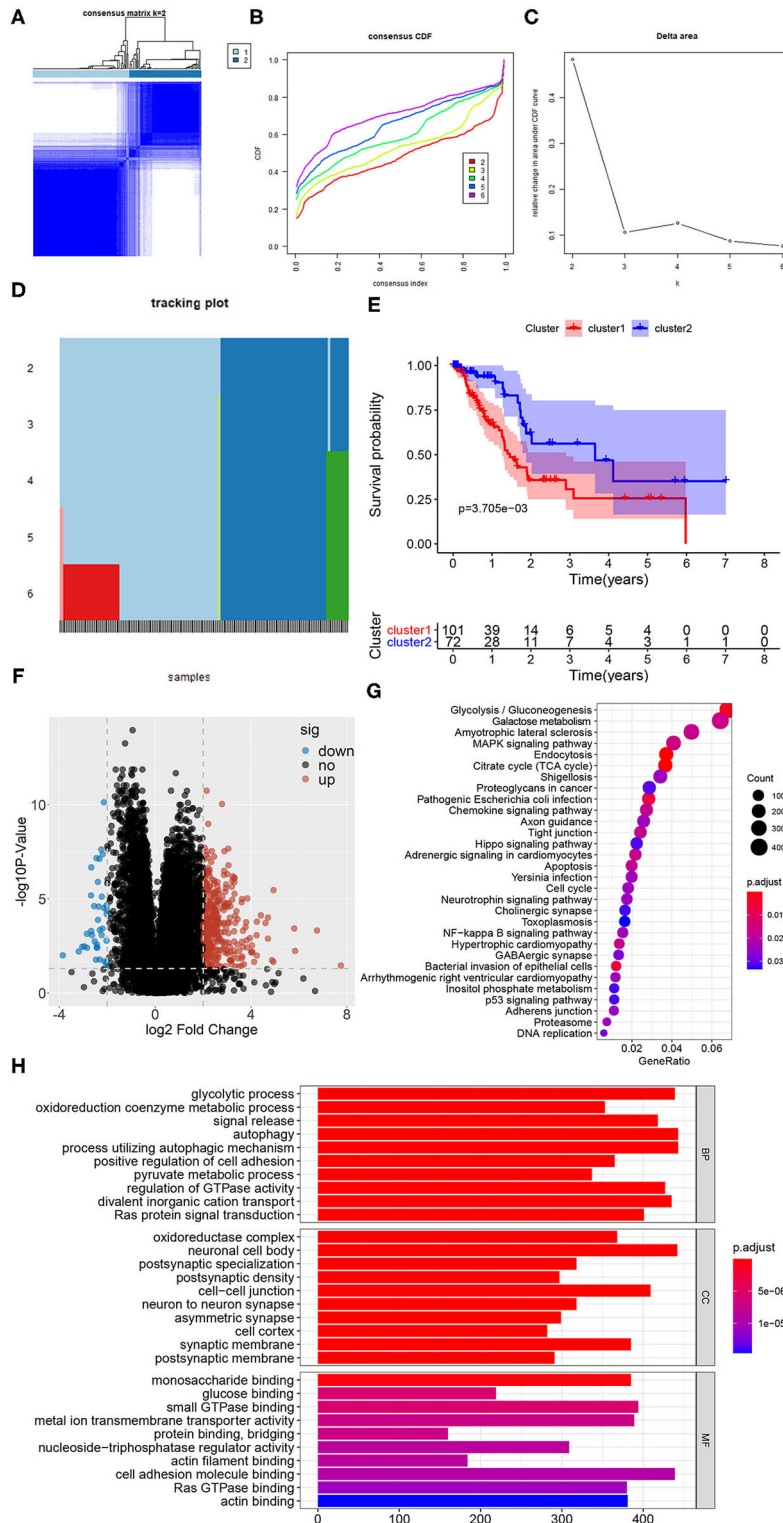
Finally, in order to study the expression level of glycolysis gene in PDAC samples at the protein level, immunohistochemical data of 10 proteins were retrieved from the HPA database, and immunohistochemical information of GPR87 was lacking. Figures 7A,B showed the staining and intensity of 10 proteins



**FIGURE 1 |** GO and KEGG pathway enrichment analysis of 319 glycolysis-related genes selected from GSEA. **(A)** The bar plot of GO pathway enrichment analysis. **(B)** The bar plot of KEGG pathway enrichment analysis (BP, biological process; CC, cell component; MF, molecular function).

in all PDAC samples. The results showed that ABCB6, CHST12, and CACNA1H were weakly stained in PDAC samples, while ALDH3B1, PGM1, MET, KIF20A, NT5E, CDK1, and

B3GNT3 were strongly stained and the protein level was higher in PDAC samples (**Figure 7C**), consistent with that at mRNA level.



**FIGURE 2 |** Identification of subclasses identification based on 319 glycolysis-related genes using NMF consensus clustering. **(A)** Consensus matrix heatmap for  $k = 2$ . **(B)** The CDF value of consensus index. **(C)** Relative change in area under CDF curve for  $k = 2-6$ . **(D)** The tracking plot for  $k = 2-6$ . **(E)** Kaplan-Meier survival analysis of PDAC patients in cluster 1 and cluster 2. **(F)** The volcano plot of the DEGs expression signature in cluster 1 and cluster 2. In the plot, “down” means down-regulated DEGs, “up” means up-regulated DEGs and “no” means the difference was not statistically significant. **(G)** The bubble plot of KEGG pathway enriched with DEGs. **(H)** The bar plot of GO pathway enriched with DEGs (BP, biological process; CC, cell component; MF, molecular function).

**TABLE 2** | Prognosis-related glycolytic genes in pancreatic cancer based on univariate Cox regression analysis.

Gene ID	HR	95% confidence interval	P-value
ALDH3B1	1.043	1.019–1.067	0.000339
DEPDC1	1.363	1.077–1.725	0.009978
CD44	1.027	1.015–1.040	1.74E-05
NUP160	1.264	1.087–1.471	0.00236
PHKA2	0.750	0.620–0.907	0.003105
PRKACA	0.896	0.835–0.962	0.002368
HMMR	1.396	1.212–1.608	3.57E-06
PGM1	1.067	1.033–1.102	8.05E-05
B4GALT1	1.012	1.005–1.019	0.001039
AURKA	1.090	1.027–1.158	0.004705
PYGL	1.068	1.032–1.106	0.00018
PYGB	1.003	1.000–1.007	0.039495
PGK1	1.007	1.002–1.012	0.009788
MET	1.046	1.033–1.060	3.98E-12
B3GAT1	0.283	0.081–0.990	0.048251
KIF20A	1.329	1.189–1.486	5.46E-07
RARS1	1.112	1.040–1.188	0.001751
CENPA	1.227	1.078–1.395	0.001907
ABCB6	0.469	0.252–0.874	0.017151
P4HA1	1.011	1.002–1.021	0.022624
LDHA	1.011	1.007–1.016	1.38E-06
NT5E	1.040	1.026–1.053	2.75E-09
CHST12	0.305	0.174–0.537	3.73E-05
GPR87	1.049	1.026–1.073	2.64E-05
NUP54	1.198	1.022–1.403	0.025428
LHX9	1.432	1.138–1.545	0.003413
EGFR	1.047	1.012–1.083	0.007303
AK4	1.076	1.009–1.146	0.024863
CDK1	1.201	1.108–1.302	9.18E-06
B3GNT3	1.015	1.006–1.024	0.00069
COPB2	1.049	1.007–1.092	0.021589
PPIA	1.013	1.002–1.025	0.018752
CACNA1H	0.820	0.728–0.924	0.001126
RPE	1.103	1.052–1.157	4.83E-05
ERO1A	1.016	1.005–1.026	0.002958

## Analysis of Immune Cell Infiltration in PDAC

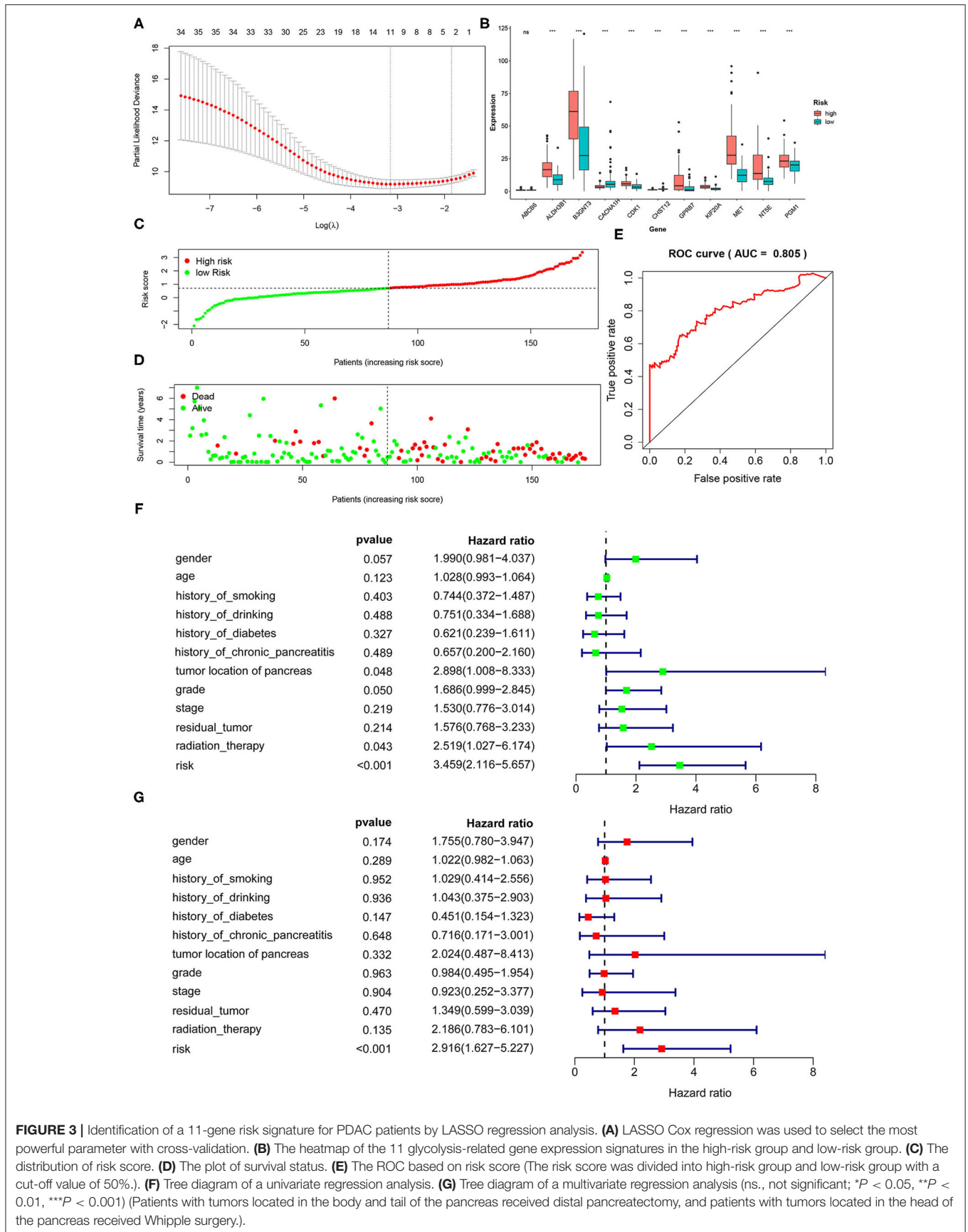
The box plot showed the infiltrating abundance of TIICs with significant differences ( $P < 0.05$ ) in PDAC samples between the high-risk and low-risk groups. It can be seen that, compared with the low-risk group, the infiltration abundance of macrophages in the high-risk samples was higher, and that of CD8+T cells was lower, which was the same as in cluster 1 (**Supplementary Figure 5A**). In addition, our analysis also found that the high-risk group had lower immune score, stromal score, and ESTIMATE score ( $P < 0.05$ ), while the tumor purity was higher ( $P < 0.05$ ) (**Supplementary Figures 5B–E**).

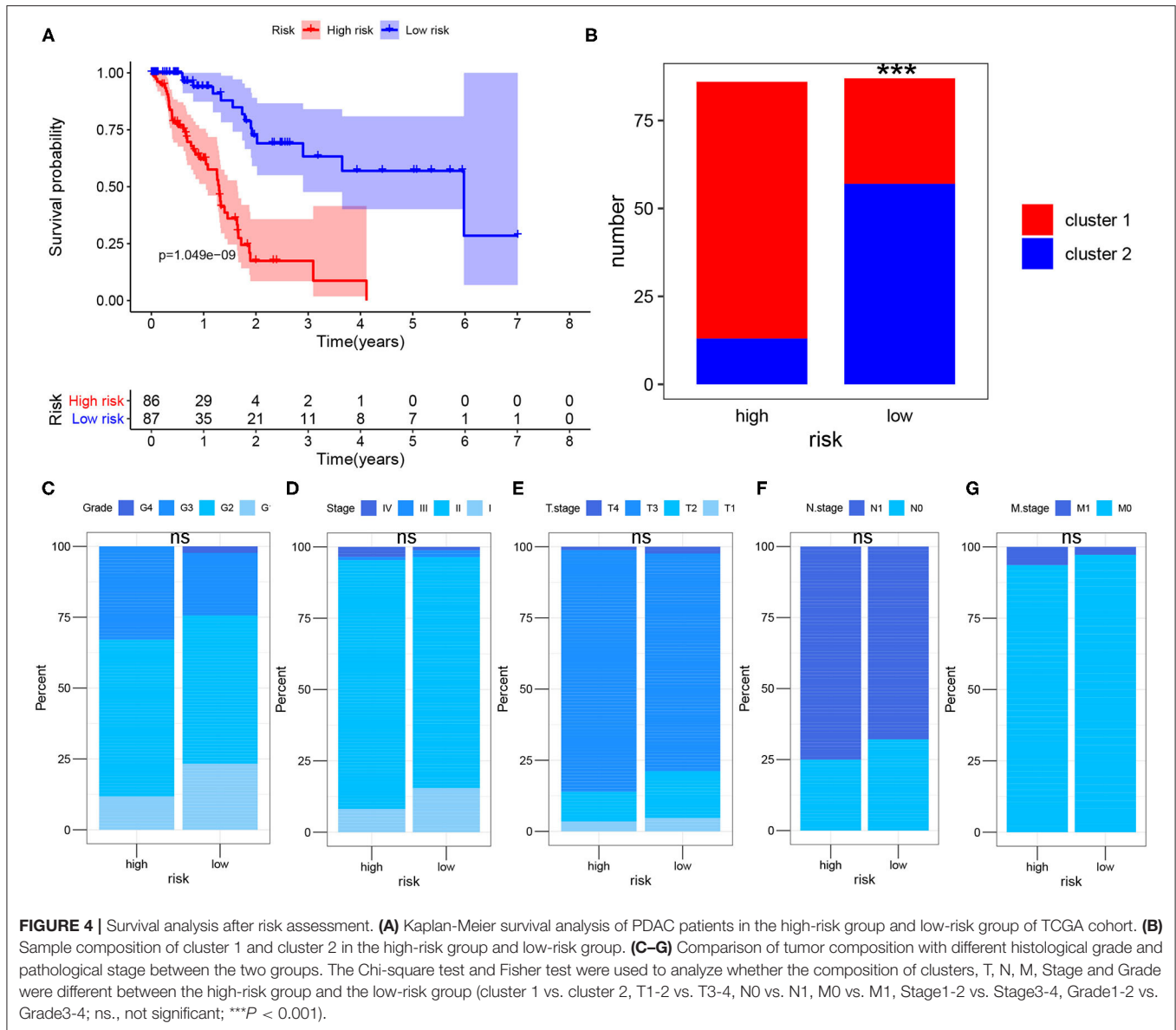
## DISCUSSION

PDAC is a common cause of cancer-related death and has a low survival rate. It is estimated that there will be 57,600 new cases and 47,050 deaths in the United States in 2020 (Siegel et al., 2020). Many methods can be used in the clinical treatment of PDAC, such as surgery, chemotherapy, radiotherapy, immunotherapy, targeted molecular therapy, etc. The individual characteristics of PDAC patients are of great significance for precision therapy, among which the genetic characteristic of tumor is a very important aspect. There have been many studies on gene signatures based on TCGA and GEO database to predict the prognosis of PDAC patients, including immune-related genes and mucin genes (Wei et al., 2019; Wu et al., 2019, 2020; Jonckheere et al., 2020). Pilar Espiau-Romera's review summarized the correspondence between the molecular subtypes and metabolic subtypes of PDAC, and the metabolic reprogramming involved in PDAC included lipid metabolism, glycolysis, and amino acid consumption. The survival analysis showed that the subtype classification of PDAC based on metabolism was more significant for clinical diagnosis and treatment, because it can not only indicate the prognosis of patients, but also reflect the invasion, drug resistance and other biological characteristics of PDAC (Espiau-Romera et al., 2020). In this study, we used the TCGA-PAAD cohort to analyze the effect of glycolysis status on the OS of PDAC patients, suggesting the potential value of glycolysis-related gene signature as prognostic markers.

First, we extracted glycolysis-related genes by searching the gene sets from GSEA, and then according to these 319 gene expression levels, the PDAC samples were divided into 2 clusters by NMF cluster analysis. It was found that cluster 1 contained more advanced samples and the prognosis of the patients was worse. Further, 38 prognostic genes were selected by K-M survival analysis and a risk scoring model based on LASSO regression analysis. ALDH3B1, PGM1, MET, KIF20A, NT5E, GPR87, CDK1, B3GNT3 were considered to play a role in promoting cancer, while CHST12, CACNA1H, and ABAC611 were thought to inhibit tumor progression, and similar expression differences were found in the protein level. Correlation analysis suggested that MET had a strong positive correlation with other glycolysis-related genes. Current research has shown that MET can establish connections between extracellular matrix and cytoplasm by binding to its ligand, hepatocyte growth factor (HGF). In cancer cells, abnormal HGF/C-MET axis promoted tumor progression by inducing PI3K/AKT, Ras/MAPK and other signaling pathways (Hervieu and Kermorgant, 2018; Zhang et al., 2018). Yan et al. demonstrated that HGF/C-met enhances the stem-like potential and glycolysis of PDAC cells by activating YAP/HIF-1 (Yan et al., 2018). ALDH3B1 was a member of the ALDH family and was generally considered to be metabolically active, with unique specificity for various aldehyde substrates (Kitamura et al., 2013). Expression pattern and clinical significance studies have found that ALDH3B1 was significantly highly expressed in lung cancer (Marchitti et al., 2010), and can be an independent prognostic factor for lung cancer (Sun et al., 2020). PGM1 was an enzyme

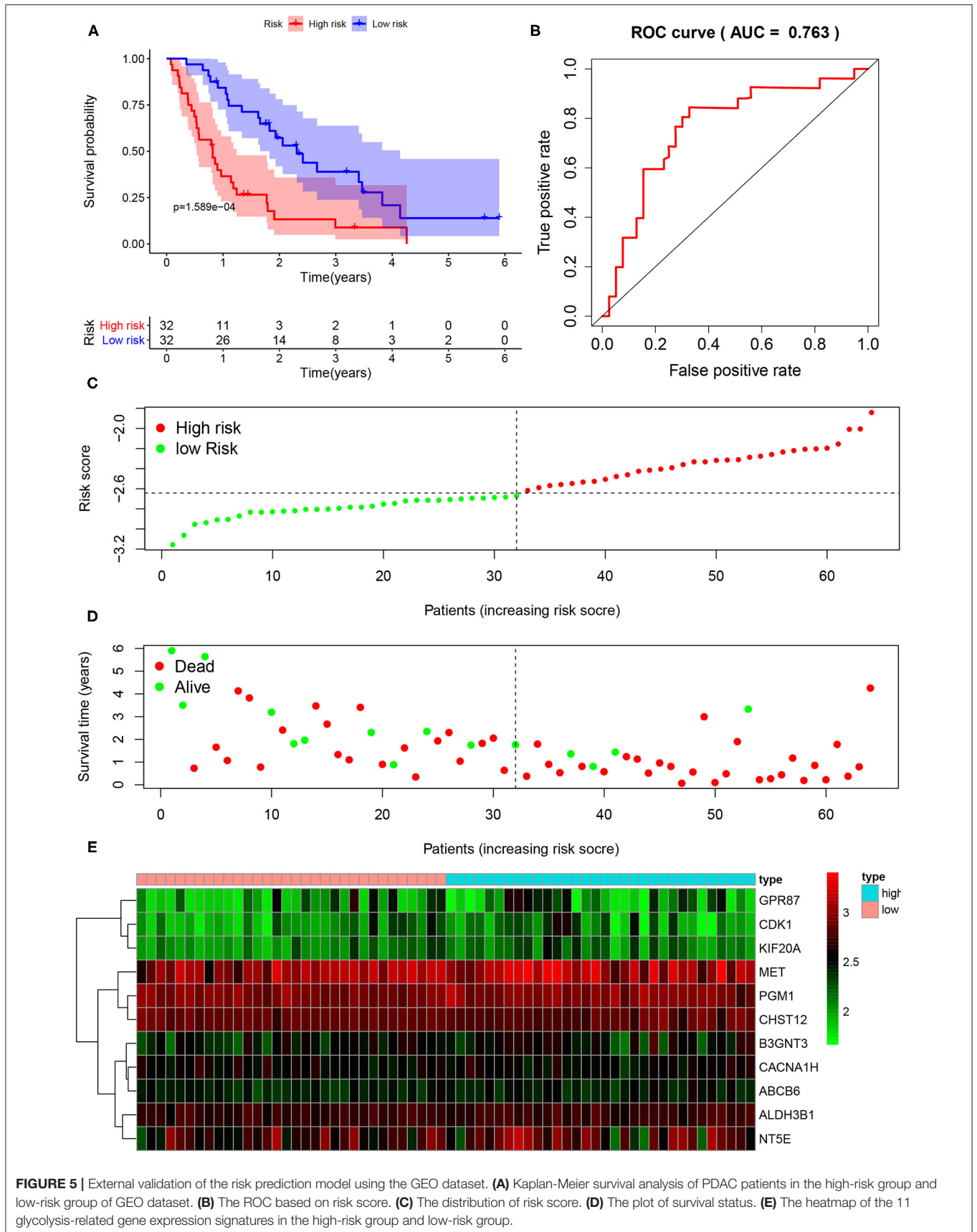




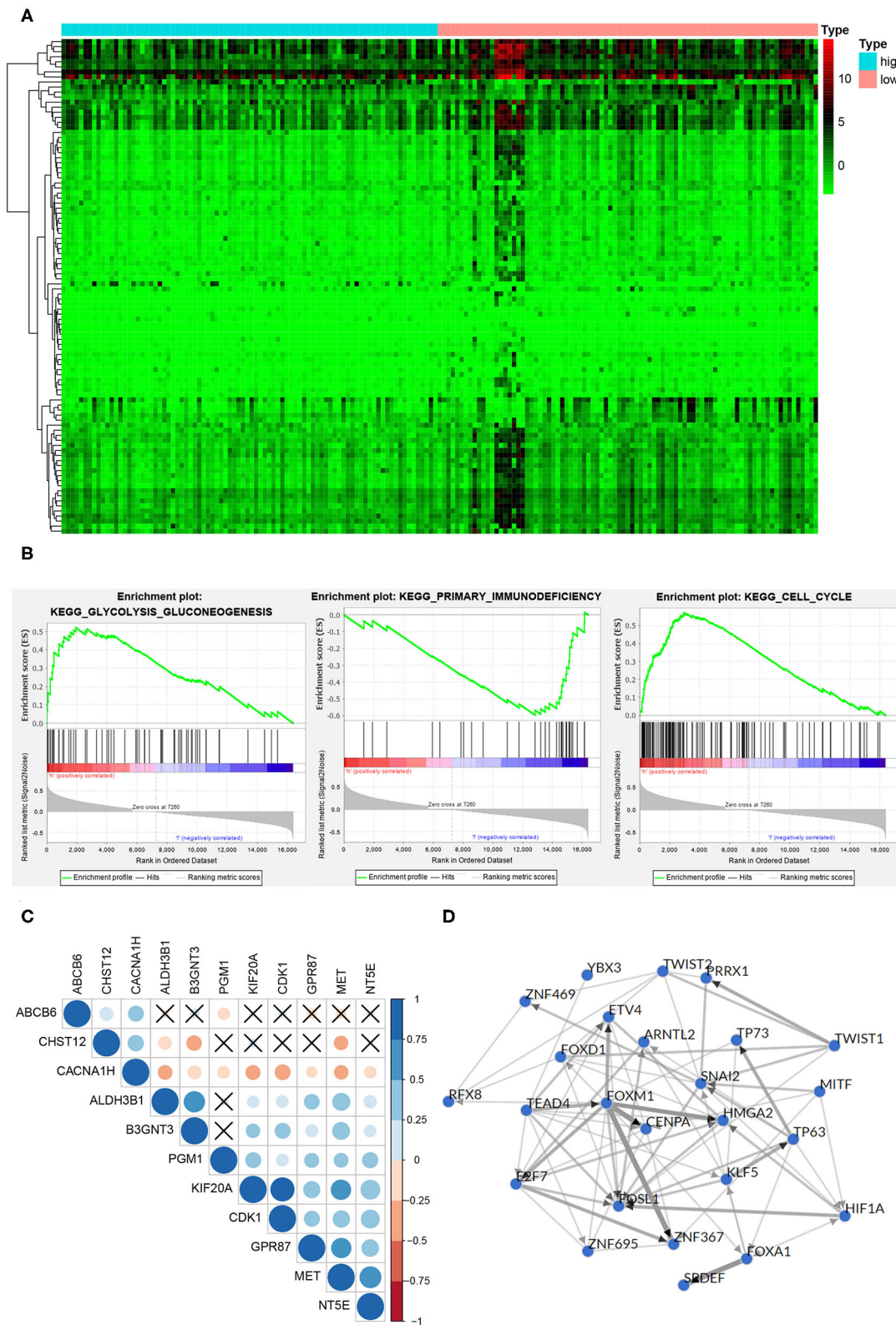


in the glycogen degradation pathway responsible for converting glucose 1 phosphate into glucose 6 phosphate. Marion et al. found that mutations in PGM1 can block its enzyme activity and prevent cancer-related fibroblast stimulation of glycogen mobilization (Curtis et al., 2019). Jung et al. found that KIF20A was up-regulated in a lactate-dependent manner to promote metastasis in the presence of excess lactic acid resulting from enhanced aerobic glycolysis in cancer (Jung et al., 2019). Yu et al. have incorporated NT5E into the glycolytic-based seven-gene signature of gastric cancer, which was closely related to the prognosis of gastric cancer patients and tumor immune infiltration (Yu et al., 2020). GPR87, CDK1, and B3GNT3 were significantly overexpressed in PDAC cells and clinical tissues, indicating poor prognosis (Wang et al., 2017; Mishra et al., 2019; Piao et al., 2019). There have also been previous studies on glycolysis-related genes in PDAC. For example, Tian et al.

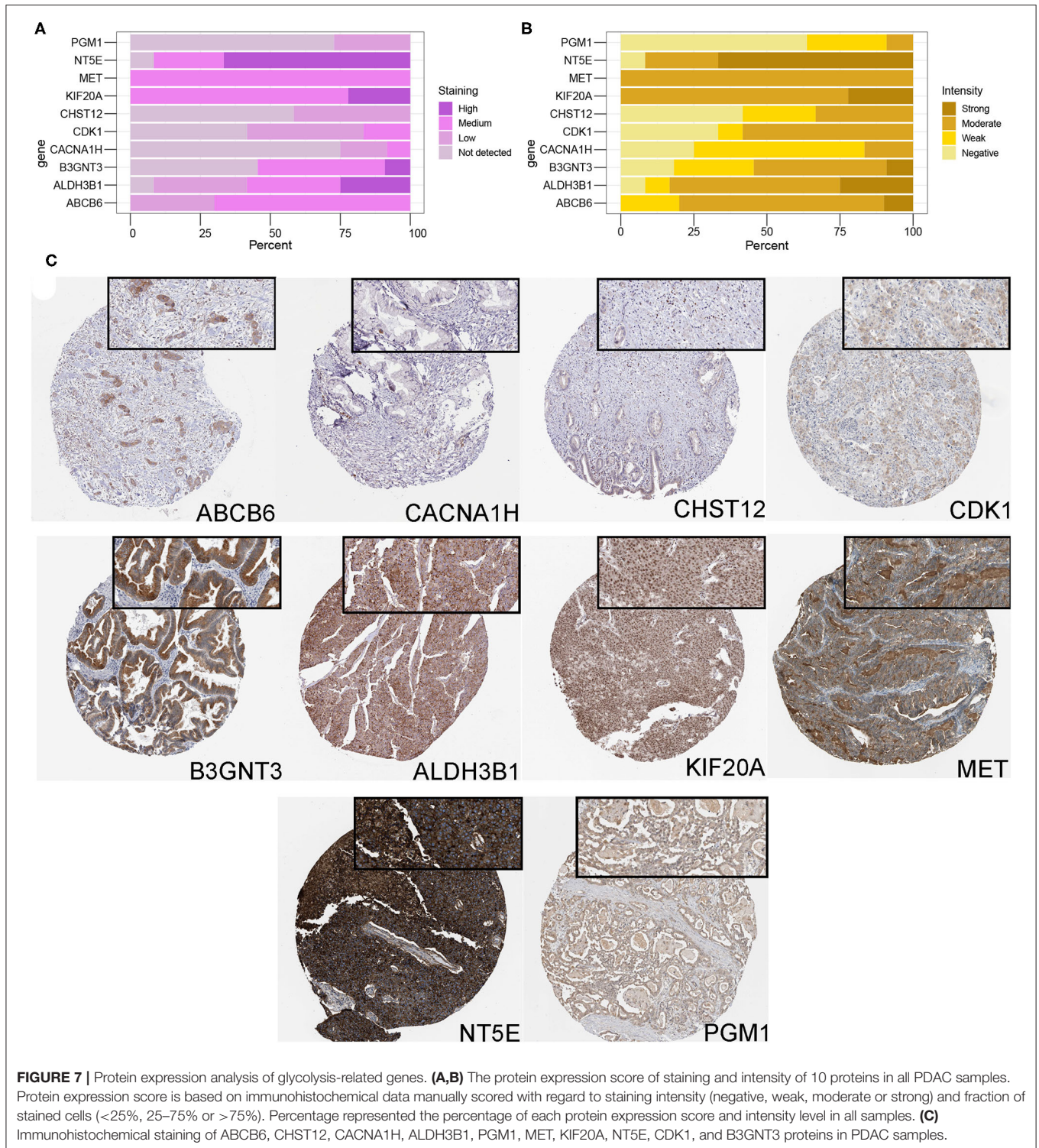
screened 13 glycolysis genes as independent prognostic factors in PDAC patients through multivariate Cox regression analysis and survival analysis, and constructed forward stepwise Cox regression model to calculate the risk score, which was  $0.700 \times \text{Met} + 0.683 \times \text{B3GNT3} + 0.662 \times \text{SPAG4}$ . Next, risk score, age, sex, tumor stage, radiotherapy, and residual tumor were included to establish a nomogram based on multivariate Cox regression analysis to predict the prognosis of PDAC. Our analysis and calculation results showed that a total of 11 glycolysis-related genes were included in the calculation of risk score, and compared with the three-gene model proposed by Tian et al., the prediction performance was better (AUC: 0.805 vs. 0.764). Moreover, we considered more factors related to PDAC pathogenesis and prognosis, such as, history of smoking, history of alcohol, history of diabetes, history of chronic pancreatitis, and tumor site (Tian et al., 2020).



**FIGURE 5 |** External validation of the risk prediction model using the GEO dataset. **(A)** Kaplan-Meier survival analysis of PDAC patients in the high-risk group and low-risk group of GEO dataset. **(B)** The ROC based on risk score. **(C)** The distribution of risk score. **(D)** The plot of survival status. **(E)** The heatmap of the 11 glycolysis-related gene expression signatures in the high-risk group and low-risk group.



**FIGURE 6 |** Function and mechanism inference of DEGs in the high-risk group and low-risk group. **(A)** The heatmap of the DEGs expression signature in the high-risk group and low-risk group. **(B)** The DEGs were enriched in glycolysis gluconeogenesis, primary immunodeficiency and cell cycle pathway. **(C)** TF networks of 11 glycolysis-related genes. Red represented a positive correlation, blue represented a negative correlation, and × represented no significant correlation. **(D)** The correlation plot of 11 glycolysis-related genes. The thickness of the line indicated the correlation.



After grouping according to the risk score, the high-risk group included more PDAC samples of late stage and high grade, and patients had a worse prognosis. External data reached the same conclusion, further demonstrating the specificity and accuracy of this genetic feature in differentiating PDAC with

different prognoses. Epidemiological studies have found that history of smoking and drinking can promote the occurrence and progression of PDAC (Ezzati et al., 2005; Hidalgo, 2010; Parkin et al., 2011; Bosetti et al., 2012). The incidence of PDAC was also found to differ by sex, possibly due to

smoking (Ilic and Ilic, 2016). In addition, chronic pancreatitis and diabetes were also found to be risk factors for PDAC (Wolpin et al., 2013; Andersen et al., 2017; Kirkegård et al., 2017). In addition, traditional AJCC TNM staging is currently recognized as the most effective prognostic tool for PDAC (Kamarajah et al., 2017). The tumor site of PDAC also had a significant difference in the prognosis of patients. Tumors located in the body or tail of pancreas often predicted a poorer prognosis, which may be caused by the more hidden onset, larger tumor size, higher risk of metastasis and difficulty of resection (van Erning et al., 2018; Tomasello et al., 2019). The present univariate and multivariate Cox regression analysis showed that the clinical pathological characteristics, such as age, gender, smoking history, drinking history, diabetes history, history of chronic pancreatitis, tumor location, histological Grade, pathological stage, residual tumor and radiation therapy had no significant effect on OS in PDAC patients, which may be due to insufficient sample size. Even so, risk score was found to be an independent prognostic factor in patients with PDAC.

In addition, the high-risk group included more cluster 1 samples, suggesting that the high-risk group had a higher level of glycolysis, leading to a poorer prognosis. GSEA analysis showed that DEGs higher expressed in high-risk group were mainly concentrated in glycolysis gluconeogenesis, primary immunodeficiency and cell cycle pathway, suggesting that PDAC in high-risk group had higher level of glycolysis, cell proliferation, and immunosuppression, which is consistent with the previous theory (Yang et al., 2020). With the discovery of abnormal glucose metabolism in cancer cells, more and more drugs targeting the glycolysis process are being developed and used in clinical trials. There were drugs targeting c-MET in clinical trials, such as onartuzumab, crizotinib, and tivantinib (Zhang et al., 2018). The usage of PGM1 inhibitors may be a therapeutic strategy to reduce the spread of metastatic abdominal cancers, such as PDAC (Curtis et al., 2019).

Currently, immunotherapy has been effective in a variety of cancers (Brahmer et al., 2012; Borghaei et al., 2015; Motzer et al., 2015; Robert et al., 2015; Sharma et al., 2016), but has not yet been converted to PDAC (Royal et al., 2010; Le et al., 2019). Most evidence indicated that there were dense stromal cells in the microenvironment of PDAC, and the immune cells were mainly myeloid suppressor cells and tumor-associated macrophages (TAM), forming an immunosuppressive environment devoid of nutrients (Dougan, 2017). Due to the abundant infiltration of bone marrow cells and the relative lack of T cells in PDAC (Liu et al., 2016), PDAC has poor response to immunotherapy treatment and poor immunotherapy effect. Raghu et al. found a survival advantage in PDAC patients with T cell infiltration in the tumor by immunohistochemical staining and multispectral imaging (Carstens et al., 2017). By analyzing the infiltration of TIICs in PDAC samples, we found that, compared with the low-risk group, the infiltration level of TIICs in the high-risk group was generally lower, M0 and M1 macrophages infiltration abundance was higher, and CD8+T lymphocytes was with low abundance, which may explain the

poor prognosis of high-risk group from the aspects of the immune mechanism. Moreover, the above characteristics of TIICs infiltration also existed in cluster 1 patients, indicating that the high-risk group had a strong similarity to cluster 1. Despite the difficulties of immunotherapy for PDAC, current research offered glimmers of hope, such as depletion or reprogramming of myeloid cells to reduce immunosuppression and fibrosis, and recruitment and enhancement of T cell response (Dougan, 2017).

Our study focused on bioinformatics to predict the diagnostic, therapeutic and prognostic value of glycolysis-related genes in PDAC, and explored the potential mechanisms by which glycolytic genes regulated of tumor cell invasion and migration. In addition, we analyzed the level of TIICs in PDAC and its relationship with prognosis. Our results indicated that the high level of glycolysis may suggest a poor prognosis in PDAC patients, which can be concluded from the comparison of signaling pathways and immune infiltration. However, we had to say that this study was a retrospective study with some limitations. Due to the low incidence of PDAC, the sample size collected by each study or center was small, which was a deficiency of this study. Therefore, we call for a larger sample size prospective study to verify the clinical application of glycolysis-related genes in personalized management of PDAC patients. We performed GO and KEGG analyses on these 319 genes, and the results showed that they were indeed enriched in glycolysis, however these genes may also be involved in regulating other processes. But complex biological processes in the human body resulted that nearly every gene could participate in multiple signaling pathways and perform multiple functions. Although we haven't found a suitable method to solve this problem now, but we will try to develop a more accurate method to screen genes for further analysis and research. In addition, this study inferred from the perspective of bioinformatics that there was a lack of experimental verification, such as the absence of resolution of glycolysis gene expression in PDAC cells, and the absence of functional studies to block or interact with glycolysis gene. In this regard, Li et al. also conducted a similar analysis and found that glycolysis level was associated with the prognosis of PDAC patients. It is worth learning that Li et al. verified that STAT3 signaling pathway was the key pathway regulating glycolysis in PDAC and there was a positive feedback between glycolysis level and STAT3 signaling activity at the cellular level (Li et al., 2021). As a continuation of future research, we will supplement it in future study.

In present study, we used the TCGA-PAAD RNA-seq data and clinical data, constructed risk prediction model based on glycolysis gene. The risk score based on this model was an independent prognostic factor for PDAC and can potentially predict the prognosis of patients. The risk prediction model was useful for verifying PDAC patients with poorer prognoses and might offer a new view for the research of individual treatment. In addition, the included glycolysis genes were significantly correlated with the invasion, cell division and cell adhesion and abundance of TIICs, and corresponding targeted drugs have been emerging.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/**Supplementary Material**, further inquiries can be directed to the corresponding author/s.

## AUTHOR CONTRIBUTIONS

WS and JZ contributed to the conception of the study. PG, XH, and YY contributed significantly to analysis and manuscript preparation. SH and YZ performed the data analyses and wrote the manuscript. LW helped perform the analysis with constructive discussions. All

authors contributed to the article and approved the submitted version.

## SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2021.639246/full#supplementary-material>

**Extended data | TCGA-PAAD dataset:**

<https://doi.org/10.6084/m9.figshare.13347536>;

GSE62452 dataset:

<https://doi.org/10.6084/m9.figshare.13347701>;

Code:

<https://github.com/songwenjing527/Glycolytic-genes-in-PDAC.git>.

## REFERENCES

- Abdel-Wahab, A. F., Mahmoud, W., and Al-Harizy, R. M. (2019). Targeting glucose metabolism to suppress cancer progression: prospective of anti-glycolytic cancer therapy. *Pharmacol. Res.* 150:104511. doi: 10.1016/j.phrs.2019.104511
- Andersen, D. K., Korc, M., Petersen, G. M., Eibl, G., Li, D., Rickels, M. R., et al. (2017). Diabetes, pancreatogenic diabetes, and pancreatic cancer. *Diabetes* 66, 1103–1110. doi: 10.2337/db16-1477
- Bailey, P., Chang, D. K., Nones, K., Johns, A. L., Patch, A. M., Gingras, M. C., et al. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 531, 47–52. doi: 10.1038/nature16965
- Borghaei, H., Paz-Ares, L., Horn, L., Spigel, D. R., Steins, M., Ready, N. E., et al. (2015). Nivolumab versus docetaxel in advanced nonsquamous non-small-cell lung cancer. *N. Engl. J. Med.* 373, 1627–1639. doi: 10.1056/NEJMoa1507643
- Bosetti, C., Bertuccio, P., Negri, E., La Vecchia, C., Zeegers, M. P., and Boffetta, P. (2012). Pancreatic cancer: overview of descriptive epidemiology. *Mol. Carcinog.* 51, 3–13. doi: 10.1002/mc.20785
- Brahmer, J. R., Tykodi, S. S., Chow, L. Q., Hwu, W. J., Topalian, S. L., Hwu, P., et al. (2012). Safety and activity of anti-PD-L1 antibody in patients with advanced cancer. *N. Engl. J. Med.* 366, 2455–2465. doi: 10.1056/NEJMoa1200694
- Carstens, J. L., Correa de Sampaio, P., Yang, D., Barua, S., Wang, H., Rao, A., et al. (2017). Spatial computation of intratumoral T cells correlates with survival of patients with pancreatic cancer. *Nat. Commun.* 8:15095. doi: 10.1038/ncomms15095
- Cheng, L., Qin, T., Ma, J., Duan, W., Xu, Q., Li, X., et al. (2019). Hypoxia-inducible factor-1 $\alpha$  mediates hyperglycemia-induced pancreatic cancer glycolysis. *Anticancer Agents Med. Chem.* 19, 1503–1512. doi: 10.2174/1871520619666190626120359
- Collisson, E. A., Sadanandam, A., Olson, P., Gibb, W. J., Truitt, M., Gu, S., et al. (2011). Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat. Med.* 17, 500–503. doi: 10.1038/nm.2344
- Curtis, M., Kenny, H. A., Ashcroft, B., Mukherjee, A., Johnson, A., Zhang, Y., et al. (2019). Fibroblasts mobilize tumor cell glycogen to promote proliferation and metastasis. *Cell Metab.* 29, 141–55.e9. doi: 10.1016/j.cmet.2018.08.007
- Dang, C. V., Le, A., and Gao, P. (2009). MYC-induced cancer cell energy metabolism and therapeutic opportunities. *Clin. Cancer Res.* 15, 6479–6483. doi: 10.1158/1078-0432.CCR-09-0889
- Deberardinis, R. J., Sayed, N., Ditsworth, D., and Thompson, C. B. (2008). Brick by brick: metabolism and tumor cell growth. *Curr. Opin. Genet. Dev.* 18, 54–61. doi: 10.1016/j.gde.2008.02.003
- Dougan, S. K. (2017). The pancreatic cancer microenvironment. *Cancer J.* 23, 321–325. doi: 10.1097/PP0.0000000000000288
- Espiau-Romera, P., Courtois, S., Parejo-Alonso, B., and Sancho, P. (2020). Molecular and metabolic subtypes correspondence for pancreatic ductal adenocarcinoma classification. *J. Clin. Med.* 9:4128. doi: 10.3390/jcm9124128
- Ezzati, M., Henley, S. J., Lopez, A. D., and Thun, M. J. (2005). Role of smoking in global and regional cancer epidemiology: current patterns and data needs. *Int. J. Cancer* 116, 963–971. doi: 10.1002/ijc.21100
- Fox, R. G., Lytle, N. K., Jaquish, D. V., Park, F. D., Ito, T., Bajaj, J., et al. (2016). Image-based detection and targeting of therapy resistance in pancreatic adenocarcinoma. *Nature* 534, 407–411. doi: 10.1038/nature17988
- Fu, Y., Liu, S., Zeng, S., and Shen, H. (2018). The critical roles of activated stellate cells-mediated paracrine signaling, metabolism and oncology in pancreatic ductal adenocarcinoma. *Mol. Cancer* 17:62. doi: 10.1186/s12943-018-0815-z
- Hervieu, A., and Kermorgant, S. (2018). The role of PI3K in met driven cancer: a recap. *Front. Mol. Biosci.* 5:86. doi: 10.3389/fmolb.2018.00086
- Hidalgo, M. (2010). Pancreatic cancer. *N. Engl. J. Med.* 362, 1605–1617. doi: 10.1056/NEJMra0901557
- Ilic, M., and Ilic, I. (2016). Epidemiology of pancreatic cancer. *World J. Gastroenterol.* 22, 9694–9705. doi: 10.3748/wjg.v22.i44.9694
- Jagust, P., de Luxán-Delgado, B., Parejo-Alonso, B., and Sancho, P. (2019). Metabolism-based therapeutic strategies targeting cancer stem cells. *Front. Pharmacol.* 10:203. doi: 10.3389/fphar.2019.00203
- Jonckheere, N., Auwerx, J., Hadj Bachir, E., Coppin, L., Boukrout, N., Vincent, A., et al. (2020). Unsupervised hierarchical clustering of pancreatic adenocarcinoma dataset from TCGA defines a mucin expression profile that impacts overall survival. *Cancers* 12:3309. doi: 10.3390/cancers12113309
- Jung, Y. D., Cho, J. H., Park, S., Kang, M., Park, S. J., Choi, D. H., et al. (2019). Lactate activates the E2F pathway to promote cell motility by up-regulating microtubule modulating genes. *Cancers* 11:274. doi: 10.3390/cancers11030274
- Kamarajah, S. K., Burns, W. R., Frankel, T. L., Cho, C. S., and Nathan, H. (2017). Validation of the American Joint Commission on Cancer (AJCC) 8th edition staging system for patients with pancreatic adenocarcinoma: a surveillance, epidemiology and end results (SEER) analysis. *Ann. Surg. Oncol.* 24, 2023–2030. doi: 10.1245/s10434-017-5810-x
- Kamisawa, T., Wood, L. D., Itoi, T., and Takaori, K. (2016). Pancreatic cancer. *Lancet* 388, 73–85. doi: 10.1016/S0140-6736(16)00141-0
- Karasinska, J. M., Topham, J. T., Kalloger, S. E., Jang, G. H., Denroche, R. E., Culibrk, L., et al. (2020). Altered gene expression along the glycolysis-cholesterol synthesis axis is associated with outcome in pancreatic cancer. *Clin. Cancer Res.* 26, 135–146. doi: 10.1158/1078-0432.CCR-19-1543
- Keenan, A. B., Torre, D., Lachmann, A., Leong, A. K., Wojciechowicz, M. L., Utti, V., et al. (2019). ChEA3: transcription factor enrichment analysis by orthogonal omics integration. *Nucleic Acids Res.* 47, W212–W224. doi: 10.1093/nar/gkz446
- Kirkegård, J., Mortensen, F. V., and Cronin-Fenton, D. (2017). Chronic pancreatitis and pancreatic cancer risk: a systematic review and meta-analysis. *Am. J. Gastroenterol.* 112, 1366–1372. doi: 10.1038/ajg.2017.218
- Kitamura, T., Naganuma, T., Abe, K., Nakahara, K., Ohno, Y., and Kihara, A. (2013). Substrate specificity, plasma membrane localization, and lipid modification of the aldehyde dehydrogenase ALDH3B1. *Biochim. Biophys. Acta* 1831, 1395–1401. doi: 10.1016/j.bbali.2013.05.007
- Le, D. T., Picozzi, V. J., Ko, A. H., Wainberg, Z. A., Kindler, H., Wang-Gillam, A., et al. (2019). Results from a phase IIb, randomized, multicenter study of GVAX pancreas and CRS-207 compared with chemotherapy in adults with previously treated metastatic pancreatic adenocarcinoma (ECLIPSE Study). *Clin. Cancer Res.* 25, 5493–5502. doi: 10.1158/1078-0432.CCR-18-2992

- Li, J., Zhu, C., Yue, P., Zheng, T., Li, Y., Wang, B., et al. (2021). Identification of glycolysis related pathways in pancreatic adenocarcinoma and liver hepatocellular carcinoma based on TCGA and GEO datasets. *Cancer Cell Int.* 21:128. doi: 10.1186/s12935-021-01809-y
- Liao, S. H., Zhao, X. Y., Han, Y. H., Zhang, J., Wang, L. S., Xia, L., et al. (2009). Proteomics-based identification of two novel direct targets of hypoxia-inducible factor-1 and their potential roles in migration/invasion of cancer cells. *Proteomics* 9, 3901–3912. doi: 10.1002/pmic.200800922
- Liberzon, A., Birger, C., Thorvaldsdóttir, H., Ghandi, M., Mesirov, J. P., and Tamayo, P. (2015). The molecular signatures database (MSigDB) hallmark gene set collection. *Cell Syst.* 1, 417–425. doi: 10.1016/j.cels.2015.12.004
- Liberzon, A., Subramanian, A., Pinchback, R., Thorvaldsdóttir, H., Tamayo, P., and Mesirov, J. P. (2011). Molecular signatures database (MSigDB) 3.0. *Bioinformatics* 27, 1739–1740. doi: 10.1093/bioinformatics/btr260
- Liu, L., Zhao, G., Wu, W., Rong, Y., Jin, D., Wang, D., et al. (2016). Low intratumoral regulatory T cells and high peritumoral CD8(+) T cells relate to long-term survival in patients with pancreatic ductal adenocarcinoma after pancreatectomy. *Cancer Immunol. Immunother.* 65, 73–82. doi: 10.1007/s00262-015-1775-4
- Marchitti, S. A., Orlicky, D. J., Brocker, C., and Vasiliou, V. (2010). Aldehyde dehydrogenase 3B1 (ALDH3B1): immunohistochemical tissue distribution and cellular-specific localization in normal and cancerous human tissues. *J. Histochem. Cytochem.* 58, 765–783. doi: 10.1369/jhc.2010.955773
- Masui, K., Tanaka, K., Akhavan, D., Babic, I., Gini, B., Matsutani, T., et al. (2013). mTOR complex 2 controls glycolytic metabolism in glioblastoma through FoxO acetylation and upregulation of c-Myc. *Cell Metab.* 18, 726–739. doi: 10.1016/j.cmet.2013.09.013
- Mishra, N. K., Southekal, S., and Guda, C. (2019). Survival analysis of multi-omics data identifies potential prognostic markers of pancreatic ductal adenocarcinoma. *Front. Genet.* 10:624. doi: 10.3389/fgene.2019.00624
- Moffitt, R. A., Marayati, R., Flate, E. L., Volmar, K. E., Loeza, S. G., Hoadley, K. A., et al. (2015). Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat. Genet.* 47, 1168–1178. doi: 10.1038/ng.3398
- Motzer, R. J., Escudier, B., McDermott, D. F., George, S., Hammers, H. J., Srinivas, S., et al. (2015). Nivolumab versus everolimus in advanced renal-cell carcinoma. *N. Engl. J. Med.* 373, 1803–1813. doi: 10.1056/NEJMoa1510665
- Newman, A. M., Steen, C. B., Liu, C. L., Gentles, A. J., Chaudhuri, A. A., Scherer, F., et al. (2019). Determining cell type abundance and expression from bulk tissues with digital cytometry. *Nat. Biotechnol.* 37, 773–782. doi: 10.1038/s41587-019-0114-2
- Parkin, D. M., Boyd, L., and Walker, L. C. (2011). The fraction of cancer attributable to lifestyle and environmental factors in the UK in 2010. *Br. J. Cancer* 105 (Suppl. 2):S77–81. doi: 10.1038/bjc.2011.489
- Pelicano, H., Martin, D. S., Xu, R. H., and Huang, P. (2006). Glycolysis inhibition for anticancer treatment. *Oncogene* 25, 4633–4646. doi: 10.1038/sj.onc.1209597
- Piao, J., Zhu, L., Sun, J., Li, N., Dong, B., Yang, Y., et al. (2019). High expression of CDK1 and BUB1 predicts poor prognosis of pancreatic ductal adenocarcinoma. *Gene* 701, 15–22. doi: 10.1016/j.gene.2019.02.081
- Puleo, F., Nicolle, R., Blum, Y., Cros, J., Marisa, L., Demetter, P., et al. (2018). Stratification of pancreatic ductal adenocarcinomas based on tumor and microenvironment features. *Gastroenterology* 155, 1999–2013.e3. doi: 10.1053/j.gastro.2018.08.033
- Rahib, L., Smith, B. D., Aizenberg, R., Rosenzweig, A. B., Fleshman, J. M., and Matrisian, L. M. (2014). Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States. *Cancer Res.* 74, 2913–2921. doi: 10.1158/0008-5472.CAN-14-0155
- Robert, C., Long, G. V., Brady, B., Dutriaux, C., Maio, M., Mortier, L., et al. (2015). Nivolumab in previously untreated melanoma without BRAF mutation. *N. Engl. J. Med.* 372, 320–330. doi: 10.1056/NEJMoa1412082
- Royal, R. E., Levy, C., Turner, K., Mathur, A., Hughes, M., Kammula, U. S., et al. (2010). Phase 2 trial of single agent Ipilimumab (anti-CTLA-4) for locally advanced or metastatic pancreatic adenocarcinoma. *J. Immunother.* 33, 828–833. doi: 10.1097/CJI.0b013e3181eec14c
- Sharma, P., Callahan, M. K., Bono, P., Kim, J., Spiliopoulou, P., Calvo, E., et al. (2016). Nivolumab monotherapy in recurrent metastatic urothelial carcinoma (CheckMate 032): a multicentre, open-label, two-stage, multi-arm, phase 1/2 trial. *Lancet Oncol.* 17, 1590–1598. doi: 10.1016/S1470-2045(16)30496-X
- Siegel, R. L., Miller, K. D., and Jemal, A. (2020). Cancer statistics, 2020. *CA Cancer J. Clin.* 70, 7–30. doi: 10.3322/caac.21590
- Singhi, A. D., Koay, E. J., Chari, S. T., and Maitra, A. (2019). Early detection of pancreatic cancer: opportunities and challenges. *Gastroenterology* 156, 2024–2040. doi: 10.1053/j.gastro.2019.01.259
- Subramanian, A., Tamayo, P., Mootha, V. K., Mukherjee, S., Ebert, B. L., Gillette, M. A., et al. (2005). Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci. U. S. A.* 102, 15545–15550. doi: 10.1073/pnas.0506580102
- Sun, H., Zhang, M., Li, L., and Huang, Z. (2020). ALDH3B1 is an independent prognostic biomarker of lung adenocarcinoma. *Technol. Cancer Res. Treat.* 19:1533033820946018. doi: 10.1177/1533033820946018
- Thul, P. J., Åkesson, L., Wiking, M., Mahdessian, D., and Geladaki, A., Ait Blal, H., et al. (2017). A subcellular map of the human proteome. *Science* 356:eaal3321. doi: 10.1126/science.aal3321
- Tian, G., Li, G., Liu, P., Wang, Z., and Li, N. (2020). Glycolysis-based genes associated with the clinical outcome of pancreatic ductal adenocarcinoma identified by the cancer genome atlas data analysis. *DNA Cell Biol.* 39, 417–427. doi: 10.1089/dna.2019.5089
- Tomasello, G., Ghidini, M., Costanzo, A., Ghidini, A., Russo, A., Barni, S., et al. (2019). Outcome of head compared to body and tail pancreatic cancer: a systematic review and meta-analysis of 93 studies. *J. Gastrointest. Oncol.* 10, 259–269. doi: 10.21037/jgo.2018.12.08
- Uhlén, M., Fagerberg, L., Hallström, B. M., Lindskog, C., Oksvold, P., Mardinoglu, A., et al. (2015). Proteomics. Tissue-based map of the human proteome. *Science* 347:1260419. doi: 10.1126/science.1260419
- Uhlen, M., Zhang, C., Lee, S., Sjöstedt, E., Fagerberg, L., Bidkhorji, G., et al. (2017). A pathology atlas of the human cancer transcriptome. *Science* 357:eaan2507. doi: 10.1126/science.aan2507
- van Erning, F. N., Mackay, T. M., van der Geest, L. G. M., Groot Koerkamp, B., van Laarhoven, H. W. M., Bonsing, B. A., et al. (2018). Association of the location of pancreatic ductal adenocarcinoma (head, body, tail) with tumor stage, treatment, and survival: a population-based analysis. *Acta Oncol.* 57, 1655–1662. doi: 10.1080/0284186X.2018.1518593
- Wang, J. J., Lei, K. F., and Han, F. (2018). Tumor microenvironment: recent advances in various cancer treatments. *Eur. Rev. Med. Pharmacol. Sci.* 22, 3855–3864. doi: 10.26355/eurrev\_201806\_15270
- Wang, L., Zhou, W., Zhong, Y., Huo, Y., Fan, P., Zhan, S., et al. (2017). Overexpression of G protein-coupled receptor GPR87 promotes pancreatic cancer aggressiveness and activates NF-κB signaling pathway. *Mol. Cancer.* 16:61. doi: 10.1186/s12943-017-0627-6
- Wei, C., Liang, Q., Li, X., Li, H., Liu, Y., Huang, X., et al. (2019). Bioinformatics profiling utilized a nine immune-related long noncoding RNA signature as a prognostic target for pancreatic cancer. *J. Cell Biochem.* 120, 14916–14927. doi: 10.1002/jcb.28754
- Weniger, M., Honselmann, K. C., and Liss, A. S. (2018). The extracellular matrix and pancreatic cancer: a complex relationship. *Cancers* 10:316. doi: 10.3390/cancers10090316
- Wolpin, B. M., Bao, Y., Qian, Z. R., Wu, C., Kraft, P., Ogino, S., et al. (2013). Hyperglycemia, insulin resistance, impaired pancreatic β-cell function, and risk of pancreatic cancer. *J. Natl. Cancer Inst.* 105, 1027–1035. doi: 10.1093/jnci/djt123
- Wu, G., Deng, Z., Jin, Z., Wang, J., Xu, B., Zeng, J., et al. (2020). Identification of prognostic immune-related genes in pancreatic adenocarcinoma and establishment of a prognostic nomogram: a bioinformatic study. *Biomed. Res. Int.* 2020:1346045. doi: 10.1155/2020/1346045
- Wu, M., Li, X., Zhang, T., Liu, Z., and Zhao, Y. (2019). Identification of a nine-gene signature and establishment of a prognostic nomogram predicting overall survival of pancreatic cancer. *Front. Oncol.* 9:996. doi: 10.3389/fonc.2019.00996
- Yan, B., Jiang, Z., Cheng, L., Chen, K., Zhou, C., Sun, L., et al. (2018). Paracrine HGF/c-MET enhances the stem cell-like potential and glycolysis of pancreatic cancer cells via activation of YAP/HIF-1α. *Exp. Cell Res.* 371, 63–71. doi: 10.1016/j.yexcr.2018.07.041
- Yang, J., Ren, B., Yang, G., Wang, H., Chen, G., You, L., et al. (2020). The enhancement of glycolysis regulates pancreatic cancer metastasis. *Cell Mol. Life Sci.* 77, 305–321. doi: 10.1007/s00018-019-03278-z
- Yang, S., He, P., Wang, J., Schetter, A., Tang, W., Funamizu, N., et al. (2016). A novel MIF signaling pathway drives the malignant character



- of pancreatic cancer by targeting NR3C2. *Cancer Res.* 76, 3838–3850. doi: 10.1158/0008-5472.CAN-15-2841
- Yoshihara, K., Shahmoradgoli, M., Martínez, E., Vegesna, R., Kim, H., Torres-Garcia, W., et al. (2013). Inferring tumour purity and stromal and immune cell admixture from expression data. *Nat. Commun.* 4:2612. doi: 10.1038/ncomms3612
- Yu, S., Hu, C., Cai, L., Du, X., Lin, F., Yu, Q., et al. (2020). Seven-gene signature based on glycolysis is closely related to the prognosis and tumor immune infiltration of patients with gastric cancer. *Front. Oncol.* 10:1778. doi: 10.3389/fonc.2020.01778
- Zhang, Q., Chen, S., Zeng, L., Chen, Y., Lian, G., Qian, C., et al. (2017). New developments in the early diagnosis of pancreatic cancer. *Expert Rev. Gastroenterol. Hepatol.* 11, 149–156. doi: 10.1080/17474124.2017.1271323
- Zhang, Y., Xia, M., Jin, K., Wang, S., Wei, H., Fan, C., et al. (2018). Function of the c-Met receptor tyrosine kinase in carcinogenesis and associated therapeutic opportunities. *Mol. Cancer.* 17:45. doi: 10.1186/s12943-018-0796-y
- Zhou, W., Liotta, L. A., and Petricoin, E. F. (2015). Cancer metabolism and mass spectrometry-based proteomics. *Cancer Lett.* 356 (2 Pt A), 176–83. doi: 10.1016/j.canlet.2013.11.003

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Song, He, Gong, Yang, Huang, Zeng, Wei and Zhang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.