Research Article

BIOMEDICAL Journal of Scientific & Technical Research

ISSN: 2574 -1241

DOI: 10.26717/BJSTR.2026.66.010290

Research on the Molecular Mechanisms and Drug Targets of Vascular Endothelial Cells in Pancreatic Cancer Using Single-Cell Sequencing Combined with Mendelian Randomization

Taoyuan Yin, Yangwei Liao, Rui Tian and Xingjun Guo*

A Division of Hepato-Pancreato-Biliary Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science & Technology, Wuhan, Hubei, 430030, China

***Corresponding author:** Xingjun Guo, A Division of Hepato-Pancreato-Biliary Surgery, Tongji Hospital, Tongji Medical College, Huazhong University of Science & Technology, Wuhan, Hubei, 430030, China

ARTICLE INFO

Received: June 15, 2026

Published: July 07, 2026

Citation: Taoyuan Yin, Yangwei Liao, Rui Tian and Xingjun Guo. Research on the Molecular Mechanisms and Drug Targets of Vascular Endothelial Cells in Pancreatic Cancer Using Single-Cell Sequencing Combined with Mendelian Randomization. Biomed J Sci & Tech Res 66(1)-2026. BJSTR. MS.ID.010290.

ABSTRACT

Pancreatic cancer is a highly malignant tumor with an extremely low five-year survival rate. Due to its insidious clinical symptoms and propensity for aggressive invasion and metastasis, identifying risk factors early is crucial for the prevention and treatment of pancreatic cancer. This study integrates single-cell sequencing and Mendelian randomization analysis by collecting pancreatic cancer-related data from the GEO and TCGA databases. The Seurat package was utilized for cell quality control and analysis, allowing for cell subpopulation annotation. Mendelian randomization was employed to assess the causal relationship between gene expression and pancreatic cancer risk, while functional analyses using CIBERSORT, GSEA, and GSVA explored the signaling pathways affected by key genes. A total of 16 distinct cell subpopulations were identified, with endothelial cells contributing the most to pancreatic cancer. Mendelian randomization analysis confirmed five risk genes associated with pancreatic cancer (CAV2, NEDD9, PDGFB, RBP7, SERPINB6), and functional analyses further revealed the potential mechanisms by which these five target genes regulate the progression of pancreatic cancer. This study elucidates the key molecular mechanisms and immune microenvironment characteristics of pancreatic cancer through single-cell sequencing and Mendelian randomization analysis, providing new insights and directions for early prevention and personalized treatment of the disease.

Keywords: Pancreatic Cancer; Endothelial Cells; Single-Cell Sequencing; Mendelian Randomization

Abbreviations: PDAC: Pancreatic Ductal Adenocarcinoma; scRNA-seq: Single-Cell RNA Sequencing; MR: Mendelian Randomization; GSEA: Gene Set Enrichment Analysis; GSVA: Gene Set Variation Analysis; GEO: Gene Expression Omnibus; SNPs: Single Nucleotide Polymorphisms; NES: Normalized Enrichment Score; AUC: Area Under the Curve

Introduction

Pancreatic ductal adenocarcinoma (PDAC) is a highly malignant tumor with a five-year survival rate of less than 10% [1]. There is a significant difference in survival time between early and advanced pancreatic cancer, making early diagnosis and intervention in key target genes crucial for improving survival outcomes [2]. However, due to the complexity of PDAC tumor microenvironment, which includes intricate cellular interactions among tumor cells, fibroblasts, endothelial cells, and various other cell types, it is challenging to identify the key molecular mechanisms that promote the progression of pancreatic cancer [3]. Single-cell RNA sequencing (scRNA-seq) allows for the detection of gene expression patterns at the single-cell level and reveals the interactions between different cell types within the tumor microenvironment [3]. The application of scRNA-seq can explore new therapeutic strategies for pancreatic cancer from the perspective of the pancreatic cancer tumor microenvironment. Although a limited number of studies on scRNA-seq in PDAC have been published, these studies carry a risk of confounding bias in the process of target gene selection and are also susceptible to reverse causation bias when making causal inferences [4].

Mendelian randomization (MR) utilizes germline genetic variation as potential instrumental variables; because these variations exhibit quasi-randomness during meiosis and are fixed at conception, MR analysis is less affected by conventional confounding factors and reverse causation bias [5,6]. By employing MR analysis on the results of scRNA-seq, it is possible to systematically identify pancreatic cancer risk-related biomarkers associated with different cell types in the tumor microenvironment. In this study, we combined scRNA-seq data of PDAC, TCGA transcriptome data, and MR methods to analyze how cell type contributes to the disease through variations in cell numbers and gene expression. This approach allowed us to systematically identify key target genes that play significant roles in the development of PDAC within critical cell subpopulations. Finally, we evaluated the biological functions of these target genes using CIBERSORT for immune infiltration analysis, as well as Gene Set Enrichment Analysis (GSEA) and Gene Set Variation Analysis (GSVA) for gene function annotation. This exploration aimed to assess the potential of these target genes as therapeutic targets within the PDAC microenvironment.

Methods

Data Acquisition

The single cell data was downloaded from the GEO (Gene Expression Omnibus) public database (GSE212966), and 10 sample data with complete single cell expression profiles were downloaded for single cell analysis. The transcriptome data including 183 pancreatic cancer samples was downloaded from TCGA database (https://portal.gdc.cancer.gov/). Exposed data came from the eQTLGen consortium database (https://www.eqtlgen.org). Participants in the outcome-related GWAS studies selected in this study were mainly people of European ancestry. Outcome summary data including 1196 cases and 475049 controls with pancreatic cancer was sourced from the ebi database (ebi-a-GCST90018893).

Quality Control

The expression profile was first read in via the Seurat package, where we filtered cells based on the total number of UMIs per cell, the number of genes expressed, and the percentage of mitochondrial reads per cell and the percentage of ribosome reads per cell. The DoubletFinder (V2.0.4) package was used to filter the double cells of each sample respectively, thus the cell quality control was completed.

Data Standardization

First, read the expression profile through the Seurat package and filter out abnormal expression samples based on the UMI count, gene number and mitochondrial gene proportion captured in each cell; then standardize, normalize, PCA, and Harmony linear dimensionality reduction of the data, and then use Use a gravel plot to observe the optimal number of pcs, and obtain the positional relationship between each cluster through UMAP nonlinear dimensionality reduction; use CellMarker and PanglaoDB databases and query the literature to find the cell types existing in the corresponding tissue and the corresponding marker genes.

Determine the Contribution of Cell Subpopulations to the Disease

We characterize the contribution of different cell subpopulations to disease by considering changes in cell number and gene expression. We identified characteristic genes for each subpopulation. Briefly, we performed bulk differential gene expression analysis and characterized the changes in this process by identifying differential genes from the disease to the control group. Then, we defined FCscore, which measures changes in the number and expression levels of characteristic genes in biological processes. Among them, FCexp(i,j) represents the expression fold change of the i-th characteristic gene in the j-th cluster, and FCscore(i,j) is the proportional fold change of the i-th characteristic gene in the j-th cluster. The FCscore of the i-th characteristic gene in the j-th cluster is defined as the square root of FCexp(i,j) divided by FCexp(i,j). Finally, the contribution of different cell subpopulations to the disease is characterized by the average FCscore of all characteristic genes in the cluster.

Mendelian Randomization Analysis

Select the SNPs associated with the locus-wide significance threshold (P < 1e-8) of each gene as potential IVs (Instrumental variables), calculate the LD (linkage disequilibrium) between SNPs, and only retain R2 < 0.001 (clumping window size=10,000kb), only keep SNPs with p2 < 5e-8. In turn, it passes through Inverse variance weighted (IVW, using meta-analysis method combined with the Wald estimate of each SNP), MR Egger (based on the assumption that the instrument strength is independent of direct effects (InSIDE)), Weighted median (the weighted median method allows up to Correctly estimate causality in 50% of cases where IVs are invalid), Weighted mode (weighted model estimation has greater ability to detect causal effects, smaller bias, and lower Type I error rate than MR-Egger regression). The causal relationship is evaluated using two statistical methods (if there is only one statistical method for the SNP in the causal relationship, only the Wald ratio is used) to obtain an overall estimate of the impact of expression of all cis- and some cross-region genes in whole blood on pancreatic cancer. Finally, the screened causal relationships are verified and analyzed through the leave-one-out method.

Heterogeneity Test Analysis

In this study, we used the Mendelian heterogeneity test to evaluate whether there is statistical heterogeneity among the single nucleotide polymorphisms (SNPs) studied. We calculated the weighted sum of squares of effect sizes and standard errors for each SNP to produce a Q value. This Q value follows a chi-square distribution with one degree of freedom for the number of SNPs studied minus one. When the p-value of the Q value is greater than 0.05, we consider that there is insufficient evidence to support the existence of heterogeneity among the SNPs effect sizes, indicating that their effects on disease risk are statistically consistent.

Immune Infiltration

The CIBERSORT algorithm was used to evaluate immune cell types in the microenvironment, which was based on the principle of support vector regression and performs deconvolution analysis on the expression matrix of immune cell subtypes. The TCGA data was analyzed using the CIBERSORT algorithm to infer the relative proportions of 22 types of immune infiltrating cells, and conduct correlation analysis on gene expression and immune cell content.

GSEA Analysis

According to the expression of key genes in patients, patients were divided into high and low expression groups, and Gene Set Enrichment Analysis (GSEA) was used to further analyze the differences in signaling pathways between the two groups. The background gene set is the version 7.0 annotated gene set downloaded from the MsigDB database. As an annotated gene set for subtype pathways, differential expression analysis of pathways between different groups is performed. Significantly enriched gene sets (adjusted p value) are analyzed based on the consistency score.

GSVA Analysis

The gene sets for Gene Set Variation Analysis (GSVA) analysis were downloaded from the Molecular signatures database, and the GSVA algorithm was used to comprehensively score each gene set to evaluate potential biological function changes in different samples.

Transcription Factor Regulatory Network

This study used the R package "RcisTarget" to predict transcription factors. All calculations performed by RcisTarget are based on motifs. The normalized enrichment score (NES) of a motif depends on the total number of motifs in the database. In addition to the motifs annotated by the source data, we inferred further annotation files based on motif similarity and gene sequence. The first step in estimating the overexpression of each motif on a gene set is to calculate the area under the curve (AUC) for each motif-motif set pair. This was performed based on recovery curve calculations of the gene set against the ordering of the motifs. The NES of each motif is calculated based on the AUC distribution of all motifs in the gene set.

Statistical Analysis

All statistical analyzes were performed using R language (version 4.3.0), and $p<0.05$ was considered statistically significant.

Results

Single Cell Analysis Identified Endothelial Cells as the Most Contributed Cell Subpopulations in Pancreatic Cancer

The single cell data of GSE212966 including a total of 10 samples was screened and filtered, and a total of 36,277 cells were retained (Supplementary Figures 1A & 1B). The 10 genes with the highest standard deviations were PRSS1, CLPS, PNLIP, CTRB1, PRSS2, CELA3A, CPA1, CPB1, CELA3B, and CTRC (Supplementary Figure 1C). The data were subjected to standardization, homogenization, PCA, and harmony analysis in sequence (Supplementary Figures 1D & 1F), and finally a total of 16 subgroups were obtained through UMAP analysis (Figure 1A). This study annotated cells by known cell markers, and 16 subpopulations were annotated to CD8+ T cells, Fibroblasts, B cells, Epithelial cells, Monocytes, CD4+ T cells, Endothelial cells, Neutrophils, Mast cells, NK cells, Neurons are the 11 cell categories (Figure 1B). In addition, we also analyzed the histogram of cell proportions corresponding to the groups (Supplementary Figure 1G) and the contribution of different cell subpopulations (Figure 1C).

We characterize the contribution of different cell subpopulations to disease by considering changes in cell number and gene expression. First, we selected the top 26 (logFC > 1) differential genes from the disease group to the control group to describe the changes in this process. Then the FCscore was defined to measure changes in the number and expression levels of characteristic genes in biological processes, and Endothelial cells, Neurons and Fibroblasts were the most contributed cells in pancreatic cancer (Figure 1D). Since the Endothelial cells contributed the most to the disease, endothelial cells were used as key cells for subsequent analysis, and genes with log2FC>1 and p_val_adj<0.05 were screened as candidate gene sets for subsequent analysis.

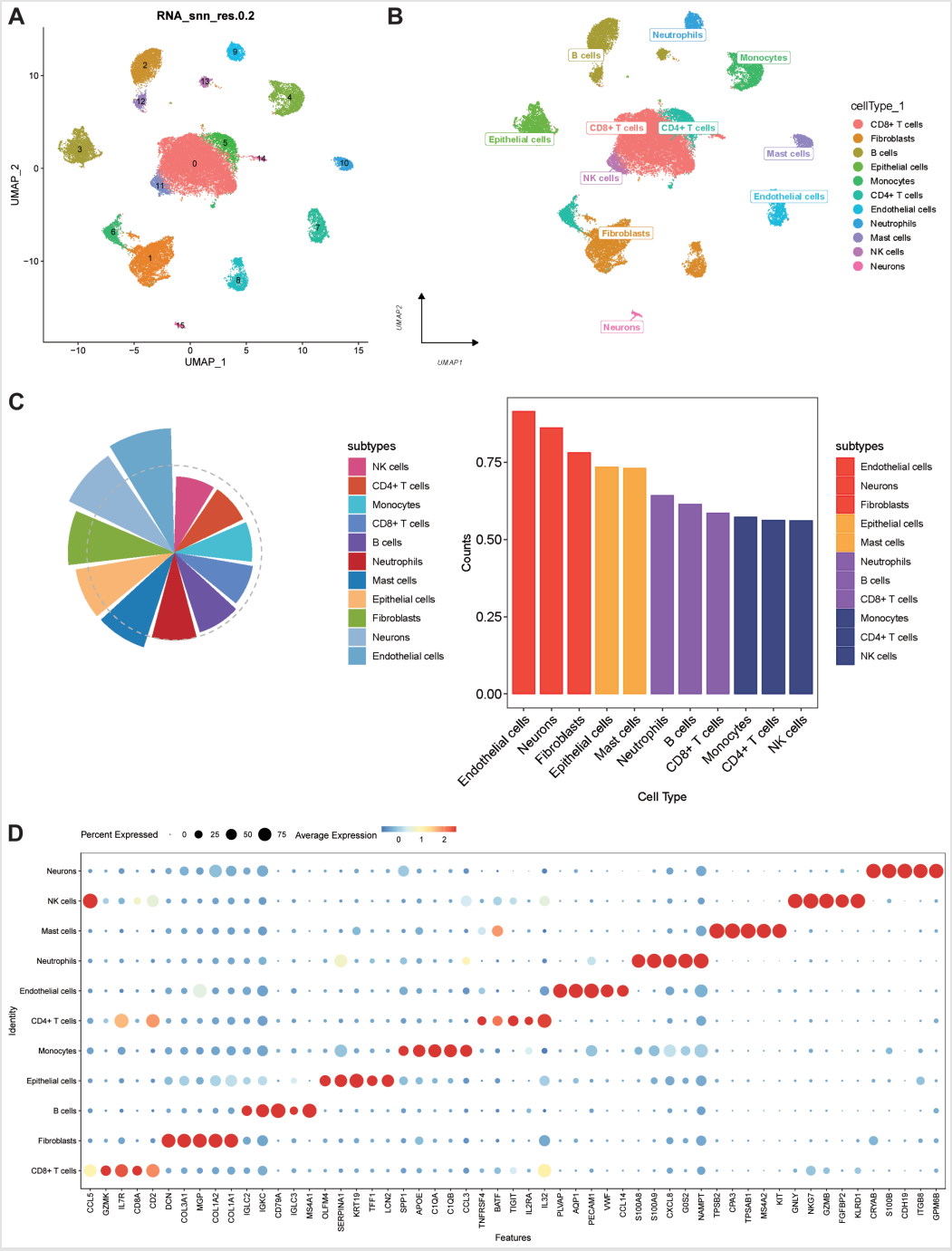


Figure 1: Single cell analysis of pancreatic cancer and contribution analysis of cell subpopulations to the disease.

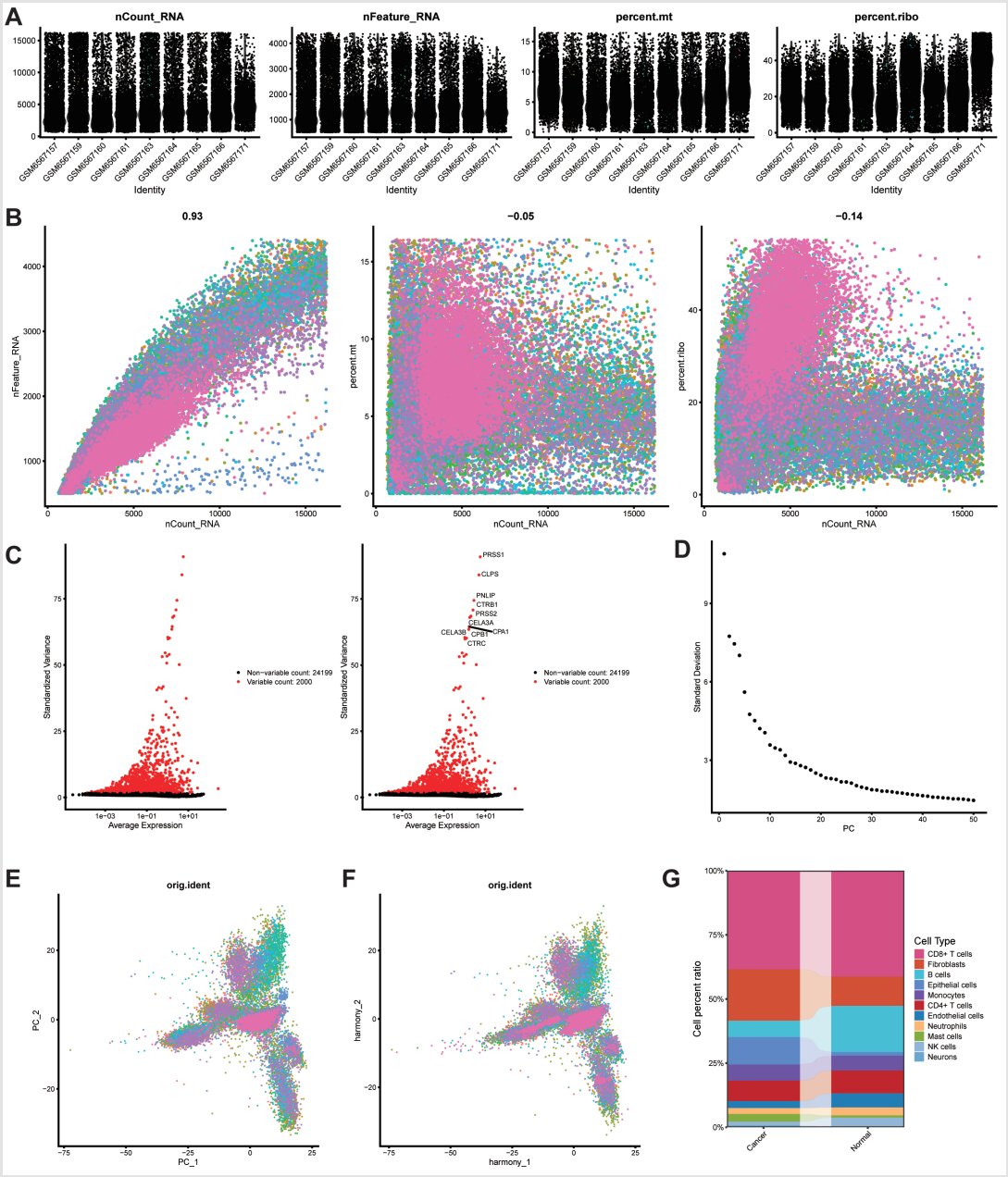
A. UMAP analysis of 36,277 cells derived from 10 samples in GSE212966. Sixteen transcriptionally distinct clusters were identified after quality control, normalization, PCA, and Harmony correction. Each color indicates a different cell cluster.

B. Cell-type annotation of the 16 clusters based on canonical marker genes. A total of 11 major cell populations were identified, including $CD8^+$ T cells, fibroblasts, B cells, epithelial cells, monocytes, $CD4^+$ T cells, endothelial cells, neutrophils, mast cells, NK cells, and neurons.

C. Pie chart and bar plot showing the contribution of different cell subpopulations.

D. Bubble plot showing the expression patterns of representative marker genes across the 11 cell types. Bubble size reflects the proportion of cells expressing each gene, and bubble color indicates expression levels. FCscore analysis integrating changes in cell number and differential gene expression (logFC > 1) between pancreatic cancer and control samples. Endothelial cells, neurons, and fibroblasts exhibited the highest FCscores, indicating their major contribution to disease progression. Endothelial cells were selected as the key cell population for subsequent analyses.

DOI: 10.26717/BJSTR.2026.66.010290



Supplementary Figure 1: Quality control, preprocessing, and preliminary characterization of single-cell data in GSE212966.

A. Quality control filtering of single-cell RNA-seq data from 10 pancreatic cancer samples. After removal of low-quality cells, doublets, and cells with high mitochondrial content, a total of 36,277 high-quality cells were retained for subsequent analysis.

B. Statistical summary of cell counts before and after filtering across all samples, showing the distribution of remaining cells included in downstream analysis.

C. The top 10 genes with the highest standard deviation across all cells, including PRSS1, CLPS, PNLIP, CTRB1, PRSS2, CELA3A, CPA1, CPB1, CELA3B, and CTRC.

D. Standardization and homogenization of gene expression matrices prior to dimensionality reduction.

E. Principal component analysis (PCA) illustrating major sources of transcriptional variability among cells.

F. Batch effect correction using Harmony, demonstrating improved integration of cells from different samples.

G. Histogram showing the proportion of annotated cell types across the 10 samples, reflecting the relative abundance of each cellular population in the dataset.

Identification of Key Target Genes by Mendelian Randomized Analysis

In order to further identify the key genes affecting pancreatic cancer among the key cell marker genes, the outcome ID was obtained through the summary statistics of 476,245 pancreatic cancer-related samples (Controls: 475,049; Cases: 1,196): ebi-a-GCST90018893. Further, through Mendelian randomization analysis, the causal relationship of 5 pairs of module genes corresponding to the positive eQTL outcome was obtained (Figures 2A-2E, IVW pval < 0.05). The corresponding genes are: CAV2, NEDD9, PDGFB, RBP7, and SERPINB6. Gene. Genes CAV2 (0.841; 0.712-0.993; P = 0.041), PDGFB (0.825; 0.698-0.976; P = 0.025), SERPINB6 (0.907; 0.837-0.982; P = 0.016) are associated with a low risk of pancreatic cancer; gene NEDD9 (1.370; 1.025-1.830; P = 0.033), RBP7 (1.180; 1.007-1.381; P = 0.040) are associated with a high risk of pancreatic cancer. Sensitivity analysis was further performed on the causal relationships of the five genes to determine their reliability. The results show that the impact on the overall error bar after excluding any one SNP is not obvious, which shows that the 5 pairs of causal relationships we selected are robust (Supplementary Figures 2A-2E). The single cell data of key genes showed that RBP7, CAV2, and PDGFB were highly expressed in endothelial cells (Figure 2F).

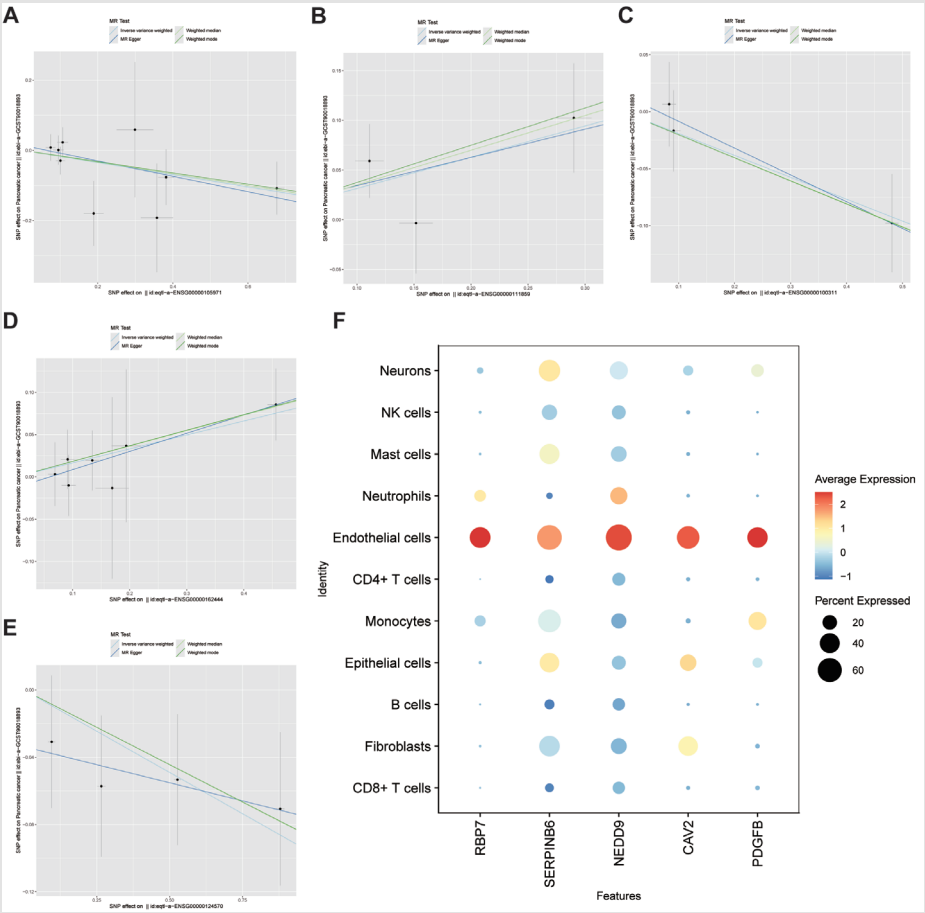


Figure 2: Mendelian randomized analysis and expressions of key target genes.

(A–E) Mendelian randomization analysis evaluating the causal associations between five endothelial cell–related genes and pancreatic cancer. Scatter plots display SNP-specific effects on gene expression and disease risk. The causal effects were estimated using the IVW method, with MR-Egger, weighted median, and weighted mode analyses performed as sensitivity analyses. (Controls: 475,049; Cases: 1,196; outcome ID: ebi-a-GCST90018893). Five gene–trait pairs (CAV2(A), NEDD9(B), PDGFB(C), RBP7(D), and SERPINB6(E)) with positive eQTL associations demonstrated significant causal relationships based on the Inverse-variance weighted (IVW) method (IVW pval < 0.05).

A. CAV2 (0.841; 0.712-0.993; P = 0.041),

B. NEDD9 (1.370; 1.025-1.830; P = 0.033),

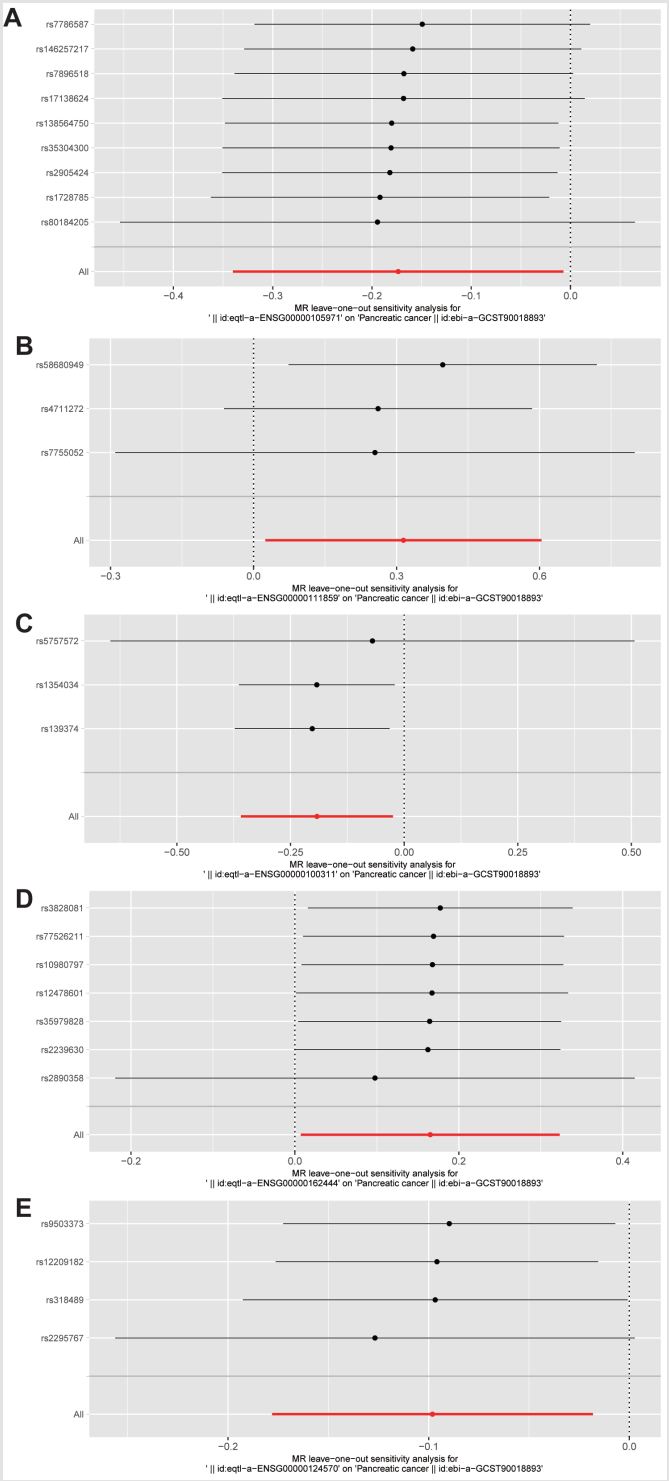
C. PDGFB (0.825; 0.698-0.976; P = 0.025),

D. RBP7 (1.180; 1.007-1.381; P = 0.040),

E. SERPINB6 (0.907; 0.837-0.982; P = 0.016).

F. Single-cell expression profiles of the five MR-identified genes, demonstrating high expression of RBP7, CAV2, and PDGFB in endothelial cells.

DOI: 10.26717/BJSTR.2026.66.010290



Supplementary Figure 2: Sensitivity analyses of Mendelian randomization results for the five key genes.

(A–E). Leave-one-out sensitivity analysis of the causal associations between each of the five MR-identified genes CAV2(A), NEDD9(B), PDGFB(C), RBP7(D), and SERPINB6(E) and pancreatic cancer. Each plot shows the effect estimates after excluding one SNP at a time. The results indicate that removal of any single SNP does not substantially affect the overall causal estimates, demonstrating the robustness and reliability of the MR findings for all five gene–trait pairs.

SERPINB6 is Functional Analysis of Upstream and Downstream of Key Target Genes

The microenvironment is mainly composed of fibroblasts, immune cells, extracellular matrix, a variety of growth factors, inflammatory factors, and special physical and chemical characteristics. The microenvironment significantly affects the diagnosis, survival outcome and clinical treatment sensitivity of the disease. We showed the distribution of immune infiltration levels and the correlation of immune cells in different forms (Supplementary Figures 3A & 3B). The research results showed that Mast cells activated, Plasma cells, and T cells CD4 memory activated were significantly high in pancreatic cancer samples (Figure 3A). We further explored the relationship between key genes and immune cells and found that CAV2 had a significant positive correlation with Macrophages M0, etc., and a significant negative correlation with NK cells activated, etc. NEDD9 has a significant positive correlation with B cells naive, etc., and a significant negative correlation with T cells regulatory (Tregs), etc. PDGFB has a significant positive correlation with Macrophages M0 and a significant negative correlation with B cells memory. RBP7 has a significant positive correlation with B cells naive, etc., and a significant negative correlation with T cells CD4 memory resting, etc. SERPINB6 has a significant positive correlation with T cells regulatory (Tregs), etc., and a significant negative correlation with T cells CD8, etc. (Figure 3B).

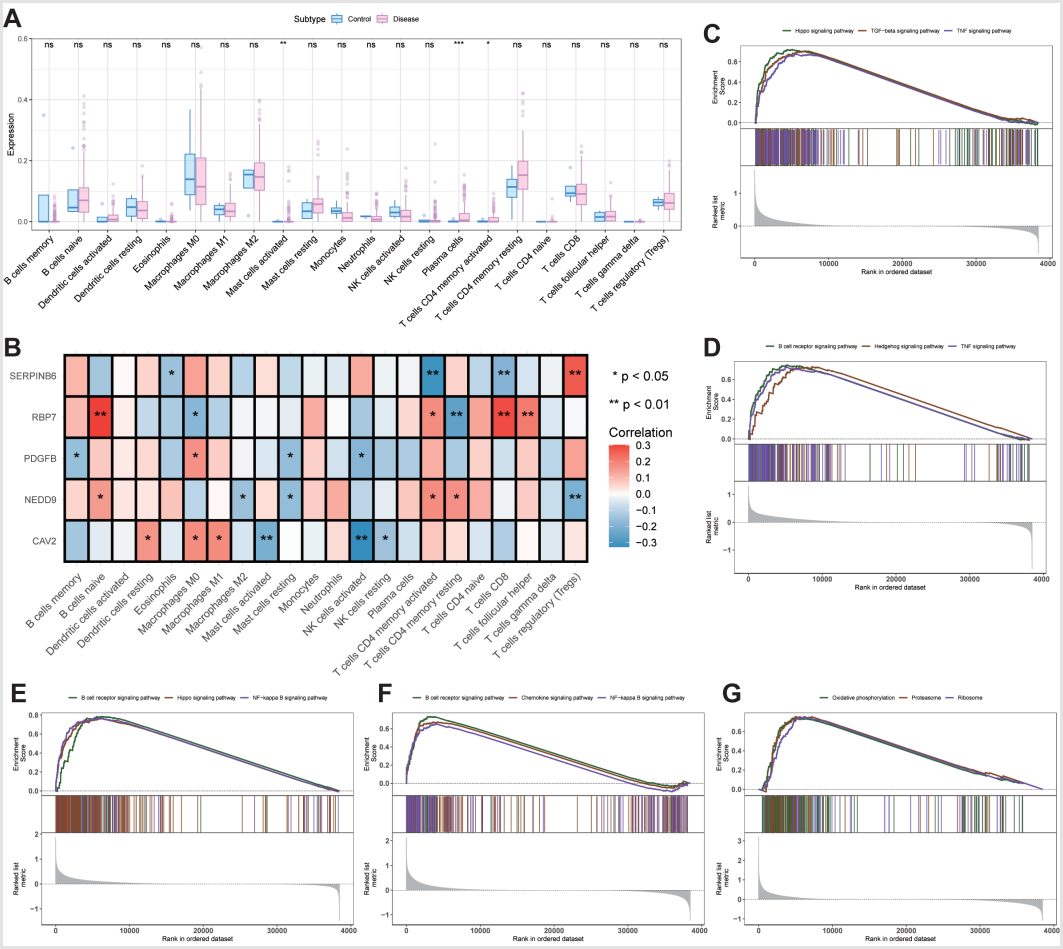


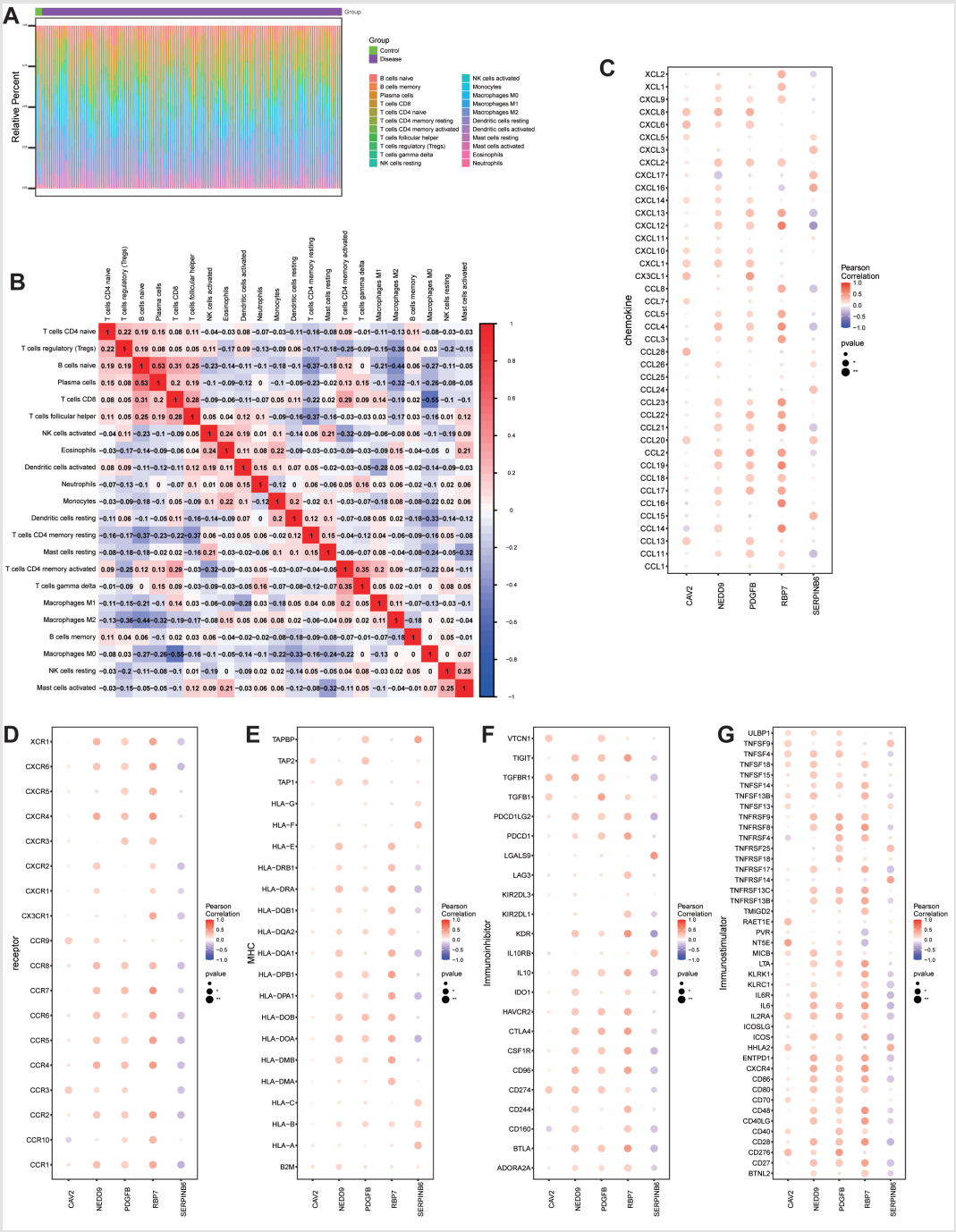
Figure 3: Immune infiltration and GSEA analyses of key target genes.

A. Immune infiltration analysis showing differences in immune cell abundance between pancreatic cancer and normal samples. Activated mast cells, plasma cells, and activated CD4 memory T cells were significantly enriched in pancreatic cancer ($P < 0.05$).

B. Correlation analysis between key target genes (CAV2, NEDD9, PDGFB, RBP7, and SERPINB6) and immune cell infiltration levels. Significant correlations were determined using Pearson correlation analysis ($P < 0.05$).

C. (C–G). Gene Set Enrichment Analysis (GSEA) showing significantly enriched signaling pathways associated with the five key genes (FDR < 0.25). CAV2 was enriched in Hippo, TGF-beta, and TNF signaling pathways. NEDD9 was enriched in TNF, B cell receptor, and Hedgehog signaling pathways. PDGFB was enriched in Hippo, NF-kappa B, and B cell receptor signaling pathways. RBP7 was enriched in chemokine, B cell receptor, and NF-kappa B signaling pathways. SERPINB6 was enriched in ribosome, oxidative phosphorylation, and proteasome pathways.The expression differences of candidate genes between two groups were assessed using the ttest, where a significance level of $P < 0.05$ was considered statistically significant. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$.

DOI: 10.26717/BJSTR.2026.66.010290



Supplementary Figure 3: Immune infiltration landscape and immune-related factor analysis in pancreatic cancer.

A. Distribution of immune cell infiltration levels between pancreatic cancer samples and normal controls, illustrating differences in immune composition within the tumor microenvironment.

B. Correlation matrix showing the relationships among different immune cell types, reflecting coordinated immune infiltration patterns in pancreatic cancer.

C. (C–G) Correlation analysis between the five key genes (CAV2, NEDD9, PDGFB, RBP7, and SERPINB6) and immune-related factors, including immunosuppressive factors, immunostimulatory factors, chemokines, and corresponding receptors. These analyses indicate that key genes are closely associated with immune regulation and play important roles in shaping the immune microenvironment.

 DOI: 10.26717/BJSTR.2026.66.010290

In addition, we analyzed the correlation between key genes and different immune factors, including immunosuppressive factors, immunostimulatory factors, chemokines, and receptors. These analyzes suggest that key genes are closely related to the level of immune cell infiltration and play an important role in the immune microenvironment (Supplementary Figures 3C-3G). Next, we studied the specific signaling pathways involved in key genes and explored the potential molecular mechanisms by which key genes affect disease progression. GSEA results show that CAV2 is enriched in the Hippo signaling pathway, TGF-beta signaling pathway, TNF signaling pathway and other signaling pathways (Figure 3C, Supplementary Figure 4A); NEDD9 is enriched in the TNF signaling pathway, B cell receptor signaling pathway, Hedgehog signaling pathway and other signaling pathways (Figure 3D, Supplementary Figure 4B); PDGFB is enriched in Hippo signaling pathway, NF-kappa B signaling pathway, B cell receptor signaling pathway and other signaling pathways (Figure 3E, Supplementary Figure 4C); RBP7 is enriched in Chemokine signaling pathway, B cell receptor signaling pathway, NF -kappa B signaling pathway and other signaling pathways (Figure 3F, Supplementary Figure 4D); SERPINB6 is enriched in Ribosome, Oxidative phosphorylation, Proteasome and other signaling pathways (Figure 3G, Supplementary Figure 4E).

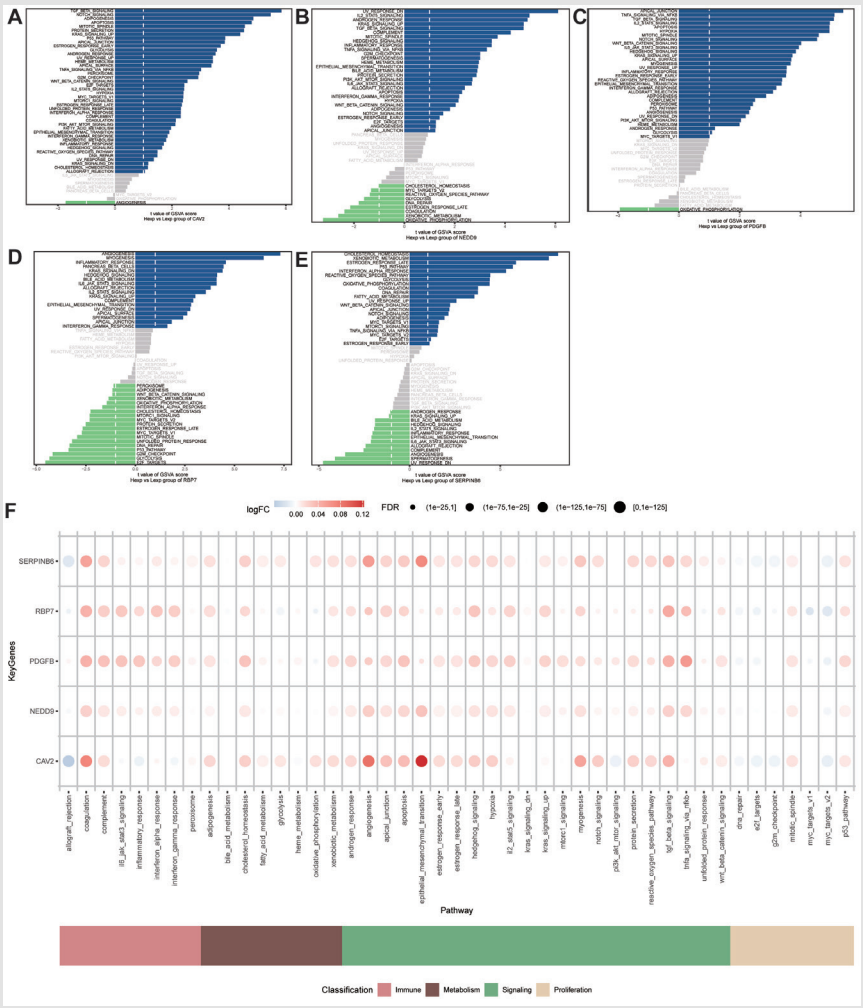


Figure 4: GSVA and immunometabolism analyses of key genes. (A-E) GSVA enrichment analysis showing signaling pathways associated with the five key genes.

A. CAV2 was enriched in TGF-beta signaling, NOTCH signaling, and other pathways.

B. NEDD9 was enriched in UV response–down, IL2–STAT5 signaling, and other pathways.

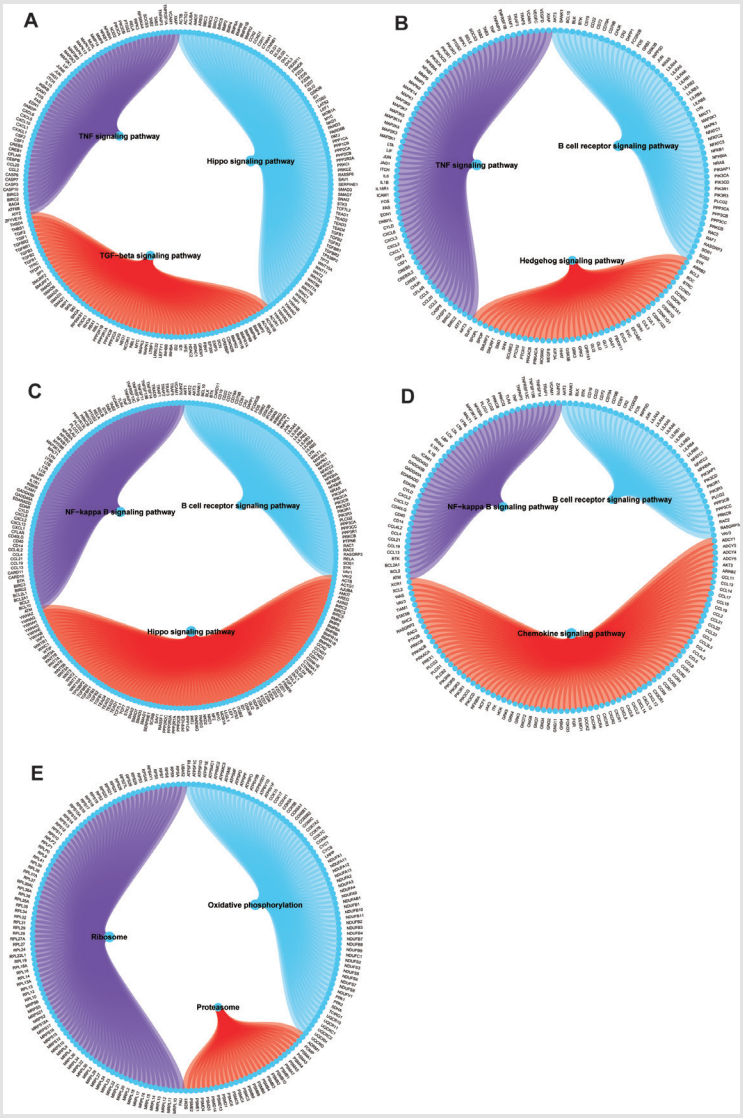
C. PDGFB was enriched in apical junction, TNFA signaling via NFKB, and other pathways.

D. RBP7 was enriched in angiogenesis, myogenesis, and other pathways.

E. SERPINB6 was enriched in cholesterol homeostasis, xenobiotic metabolism, and other pathways. These results suggest that key genes may influence pancreatic cancer progression through multiple regulatory pathways.

F. AUCell analysis of immunometabolism-related gene sets in single-cell data, with bubble plots showing the differences in expression levels of key genes across immunometabolic pathways. The expression differences of candidate genes between two groups were assessed using the ttest, where a significance level of $P < 0.05$ was considered statistically significant. *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$; ****$P < 0.0001$.

DOI: 10.26717/BJSTR.2026.66.010290



Supplementary Figure 4: GSEA enrichment analysis of signaling pathways associated with key target genes.
(A–E). GSEA results showing significantly enriched signaling pathways associated with the five key genes.

A. CAV2 was enriched in Hippo signaling pathway, TGF-beta signaling pathway, TNF signaling pathway, and others.
B. NEDD9 was enriched in TNF signaling pathway, B cell receptor signaling pathway, Hedgehog signaling pathway, and others.
C. PDGFB was enriched in Hippo signaling pathway, NF-kappa B signaling pathway, B cell receptor signaling pathway, and others.
D. RBP7 was enriched in chemokine signaling pathway, B cell receptor signaling pathway, NF-kappa B signaling pathway, and others.
E. SERPINB6 was enriched in ribosome, oxidative phosphorylation, proteasome, and other pathways.

In addition, GSVA analysis shows that CAV2 is enriched in signaling pathways such as TGF-BETA signaling and NOTCH signaling (Figure 4A); NEDD9 is enriched in signaling pathways such as UV response-down and IL2-STAT5 signaling (Figure 4B); PDGFB is enriched in signaling pathways such as apical junction and TNFA signaling via NFKB (Figure 4C); RBP7 is enriched in angiogenesis, myogenesis and other signaling pathways (Figure 4D); enriched in cholesterol homeostasis, xenobiotic metabolism and other signaling pathways (Figure 4E). This suggests that key genes may affect the progression of pancreatic cancer through these pathways. We used AUCell to quantitatively score immunometabolism-related pathway genes in single cells, and used bubble charts to display the expression differences of key genes in immunometabolism-related pathways (Figure 4F).

Relationship Between Key Target Genes and Disease-Related Genes

We took key genes as the gene set for this analysis and found that they are regulated by common mechanisms such as multiple transcription factors (Figure 5A). Therefore, enrichment analysis of these transcription factors was performed using cumulative recovery curves. Motif-TF annotation and selection analysis of important genes showed that the Motif motif with the highest normalized enrichment score (NES: 4.57) is cisbp_M5084 (Figure 5B). We obtained tumor regulation-related genes through the GeneCards database (https://www.genecards.org/). This study analyzed the expression levels of 20 genes that ranked high in Relevance*score and were expressed in the transcriptome. It was found that the expression levels of key genes were significantly correlated with the expression levels of tumor regulatory genes, among which CAV2 and EGFR were significantly positively correlated. (cor = 0.629); SERPINB6 and APC were significantly negatively correlated (cor = -0.437) (Figures 5C-5E).

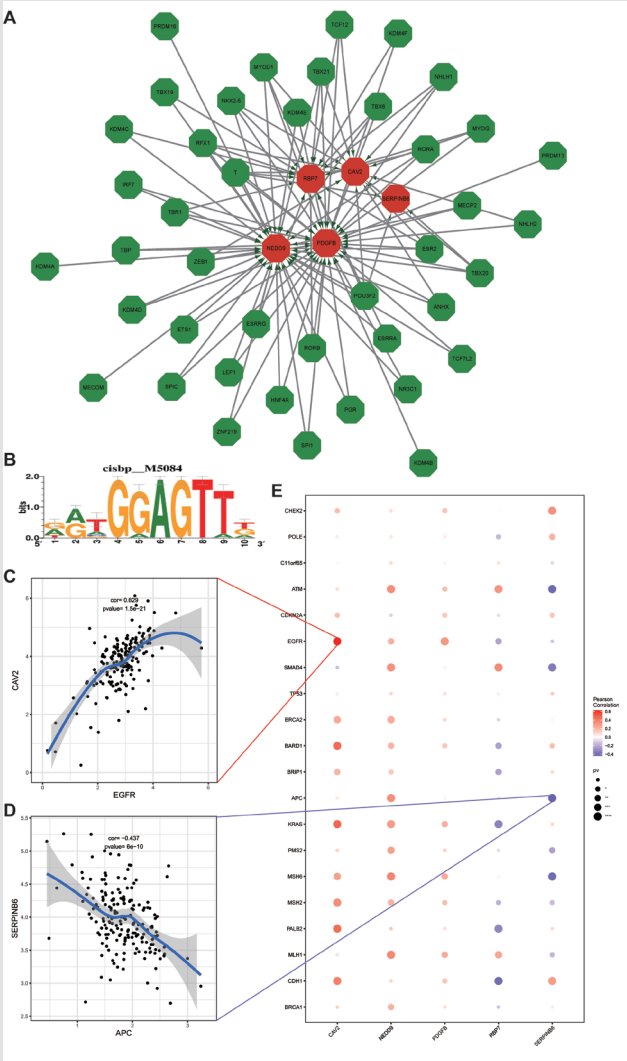


Figure 5: Transcription factor regulatory network analyses between key target genes and disease-related genes.

A. Regulatory network analysis showing that the five key genes are commonly regulated by multiple transcription factors, indicating shared upstream regulatory mechanisms.

B. Cumulative recovery curve–based transcription factor enrichment analysis. Motif–TF annotation identified cisbp_M5084 as the most significantly enriched motif with the highest normalized enrichment score (NES = 4.57). (C–E) Correlation analysis between the expression levels of key target genes and tumor regulation–related genes obtained from the GeneCards database. Twenty tumor-related genes with high Relevance*score and detectable expression in the transcriptome were analyzed.

C. Strong positive correlation between CAV2 and EGFR, cor = 0.629, P value <0.0001.

D. Negative correlation between SERPINB6 and APC, cor = −0.437, P value <0.0001.

Discussion

In this study, we analyzed gene expression in a total of 36,277 cells from PDAC tissues, identifying 11 distinct cell types, including cancer cells, fibroblasts, and endothelial cells. Through an analysis of disease contribution, we examined the number of feature genes and changes in expression levels across biological processes, revealing that endothelial cells had the highest contribution to the disease. Subsequently, through MR analysis, we selected five genes (CAV2, NEDD9, PDGFB, RBP7, SERPINB6) as the most significant genes influencing PDAC in endothelial cells. Notably, CAV2, PDGFB, and SERPINB6 were associated with a low risk of PDAC, while NEDD9 and RBP7 were linked to a high risk. Finally, we conducted immune infiltration analysis, GSEA, GSVA, transcriptional analysis, and disease gene correlation analysis to further elucidate the impact of these key genes on PDAC progression. This comprehensive approach highlights the potential of these genes as critical players in the tumor microenvironment and their relevance in therapeutic strategies for PDAC. Due to the unique tumor microenvironment, the vascular system in PDAC exhibits characteristics distinct from those found in normal tissues, promoting immune evasion and chemotherapy resistance in PDAC [7].

The tumor stroma of PDAC contains a large number of aberrant blood vessels induced by VEGF, with loosely arranged vascular endothelial cells [8]. Moreover, due to fibrotic proliferation within the stroma, the intratumoral vascular pressure is higher than that in the surrounding tissue, contributing to the formation of a hypoxic environment within PDAC and hindering the recruitment of immune cells [8]. In addition, the abnormal endothelial cells in PDAC also express receptors such as PD-L1 and FASL, which induce T-cell apoptosis, thereby promoting immune tolerance. In summary, vascular endothelial cells in PDAC may play a crucial role in the formation of an immune-tolerant microenvironment [8,9]. Targeting these endothelial cells could be an important strategy for improving therapeutic outcomes in PDAC, highlighting their potential as significant therapeutic targets in combating the challenges associated with this aggressive cancer. Prior studies suggest that the target genes identified in this research may be involved in tumor progression through vascular endothelial cells. NEDD9, also known as CASL and HEF1, encodes a multidomain scaffold protein that is known to participate in various processes, including cell proliferation, DNA damage repair, and cell migration [10].

Research has reported that NEDD9 can influence the invasion and migration of tumor cells in breast cancer and oral squamous cell carcinoma by affecting the activity of the MMP family of proteases and the phosphorylation of kinases like FAK [10-12]. In vascular endothelial cells, hypoxia or inflammatory damage can induce the expression of NEDD9 on the cell surface, where it interacts with platelet P-selectin [13]. Additionally, NEDD9 is involved in the synthesis of collagen COL3A1 in fibroblasts via exosome-mediated signaling, participating in vascular fibrosis and the development of chronic pulmonary hypertension [14]. RBP7, a member of the CRBP family, has an affinity for all-trans-retinoic [15]. It has been reported to be related to the sensitivity of tumors to chemotherapy agents such as 5-FU and tamoxifen [15,16]. RBP7 may also regulate fatty acid metabolism in breast cancer cells by inhibiting the AKT/SREBP1 pathway, thereby playing a role in tumor immune responses [17]. CAV2 encodes caveolin-2, an important component of the cytoplasmic membrane, which has been associated with poor prognosis in breast cancer, esophageal cancer, and pancreatic cancer [18]. In vascular endothelial cells, CAV2 is primarily involved in immune inflammatory responses, with studies indicating its role in leukocyte adhesion to endothelial cells during ischemia-reperfusion injury and aging processes [19,20].

PDGFB, a member of the platelet-derived growth factor family, is a key regulator of mesenchymal cell differentiation and is predominantly expressed in vascular endothelial cells, playing a significant role in tumor angiogenesis [21]. While the abundance of PDGFB in the pancreatic cancer microenvironment promotes tumor growth, recent research shows that completely eliminating PDGFB can paradoxically induce hypoxia and enhance epithelial-mesenchymal transition, facilitating the metastasis of pancreatic cancer [22]. Finally, SERPINB6 functions to inhibit serine protease activity within cells, primarily expressed in endothelial cells, epithelial cells, and certain myeloid cells. It can inhibit activities like trypsin-like protease activity, consequently affecting cell autophagy and apoptosis [23]. Overall, these genes present promising avenues for understanding pancreatic cancer's progression and potential therapeutic targets for intervention. Our study has several notable advantages. First, by using scRNA-seq data from PDAC, we quantitatively assessed the contributions of different cell subpopulations to disease progression based on changes in cell numbers and gene expression. This approach provides a more intuitive comparison of the importance of various cell types in PDAC. Second, we employed MR methods to minimize the confounding biases and reverse causation biases that can occur in the gene selection process.

This strongly suggests that these target genes are risk factors for PDAC rather than changes in gene expression driven by PDAC. Third, we applied various functional analysis methods, including immune infiltration analysis, GSEA, GSVA, and transcription factor regulatory network analysis, to provide evidence from different angles supporting the role of the target genes in promoting PDAC progression. These analyses also suggest the molecular mechanisms by which these proteins are involved in disease progression. To our knowledge, this may be the first MR analysis conducted on single-cell sequencing data related to PDAC. Finally, it is essential to note that there are limitations to this study. Although the MR analysis provides insights into causal relationships and reduces the risk of reverse causation bias, experimental studies are still needed for validation. Furthermore, since this study is based on scRNA-seq data, it cannot assess the impact of protein levels and post-translational modifications on tumor progression, highlighting the need for further research to gain a more in-depth un-

derstanding of the mechanisms of PDAC progression. In summary, despite some limitations, our research's innovation and methodological rigor lay a solid foundation for further exploration of the molecular mechanisms and potential therapeutic targets for PDAC.

Conclusion

Our study identified five endothelial cell genes associated with the risk of PDAC through MR analysis, providing new insights into the progression of PDAC and presenting promising targets for the development of drugs that target endothelial cells in PDAC. Further experimental and clinical research is needed to evaluate the efficacy of these candidate genes.

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.

Authors' Contributions

Conceptualization: TYY, RT and XJG. Methodology and analysis: TYY and YWL. Manuscript writing: TYY and YWL. Review and editing: all authors. All authors have read and agreed to the published version of the manuscript.

Ethics, Consent to Participate, and Consent to Publish Declarations

Not applicable.

Funding

This study was supported by grants from Hubei Provincial Natural Science Foundation of China (2023AFB717 to XG).

Data Availability Declaration

All data analyzed in this study are publicly available. Single-cell RNA sequencing data were obtained from the Gene Expression Omnibus (GEO) under accession number GSE212966. Bulk transcriptome data of pancreatic cancer were downloaded from The Cancer Genome Atlas (TCGA). eQTL data were obtained from the eQTLGen Consortium, and pancreatic cancer GWAS summary statistics were retrieved from the EBI GWAS Catalog (accession ID: ebi-a-GCST90018893). No new data were generated in this study.

References

1. Siegel RL, Miller KD, Fuchs HE, Jemal A (2021) Cancer Statistics. CA Cancer J Clin 71(1): 7-33.
2. Siegel RL, Giaquinto AN, Jemal A (2024) Cancer statistics. CA Cancer J Clin 74(1): 12-49.
3. Bernard V, Semaan A, Huang J, San Lucas FA, Mulu FC, et al. (2019) Single-Cell Transcriptomics of Pancreatic Cancer Precursors Demonstrates Epithelial and Microenvironmental Heterogeneity as an Early Event in Neoplastic Progression. Clin Cancer Res 25(7): 2194-2205.
4. Moncada R, Barkley D, Wagner F, Chiodin M, Devlin JC, et al. (2020) Integrating microarray-based spatial transcriptomics and single-cell RNA-seq reveals tissue architecture in pancreatic ductal adenocarcinomas. Nat Biotechnol 38(3): 333-342.
5. Richmond RC, Davey Smith G (2022) Mendelian Randomization: Concepts and Scope. Cold Spring Harb Perspect Med 12(1).
6. Davies NM, Holmes MV, Davey Smith G (2018) Reading Mendelian randomisation studies: a guide, glossary, and checklist for clinicians. Bmj 362: k601.
7. Liu Z, Wang Y, Huang Y, Kim BYS, Shan H, et al. (2019) Tumor Vasculatures: A New Target for Cancer Immunotherapy. Trends Pharmacol Sci 40(9): 613-623.
8. Jain RK (2014) Antiangiogenesis strategies revisited: from starving tumors to alleviating hypoxia. Cancer Cell 26(5): 605-622.
9. Motz GT, Santoro SP, Wang LP, Garrabrant T, Lastra RR, et al. (2014) Tumor endothelium FasL establishes a selective immune barrier promoting tolerance in tumors. Nat Med 20(6): 607-615.
10. Hu Z, Wei F, Su Y, Wang Y, Shen Y, et al. (2023) Histone deacetylase inhibitors promote breast cancer metastasis by elevating NEDD9 expression. Signal Transduct Target Ther 8(1): 11.
11. McLaughlin SL, Ice RJ, Rajulapati A, Kozyulina PY, Livengood RH, et al. (2014) NEDD9 depletion leads to MMP14 inactivation by TIMP2 and prevents invasion and metastasis. Mol Cancer Res 12(1): 69-81.
12. Grauzam S, Brock AM, Holmes CO, Tiedeken JA, Boniface SG, et al. (2018) NEDD9 stimulated MMP9 secretion is required for invadopodia formation in oral squamous cell carcinoma. Oncotarget 9(39): 25503-25516.
13. Alba GA, Samokhin AO, Wang RS, Zhang YY, Wertheim BM, et al. (2021) NEDD9 Is a Novel and Modifiable Mediator of Platelet-Endothelial Adhesion in the Pulmonary Circulation. Am J Respir Crit Care Med 203(12): 1533-1545.
14. Samokhin AO, Stephens T, Wertheim BM, Wang RS, Vargas SO, et al. (2018) NEDD9 targets COL3A1 to promote endothelial fibrosis and pulmonary arterial hypertension. Sci Transl Med 10(445).
15. Napoli JL (2017) Cellular retinoid binding-proteins, CRBP, CRABP, FABP5: Effects on retinoid metabolism, function and related diseases. Pharmacol Ther 173: 19-33.
16. Huang X, Ke K, Jin W, Zhu Q, Zhu Q, et al. (2022) Identification of Genes Related to 5-Fluorouracil Based Chemotherapy for Colorectal Cancer. Front Immunol 13: 887048.
17. Yu Y, Xu Z, Zhou H, Xu R, Xu J, et al. (2024) RBP7 functions as a tumor suppressor in HR + breast cancer by inhibiting the AKT/SREBP1 pathway and reducing fatty acid. Cancer Cell Int 24(1): 118.
18. Zhu Y, Tian J, Peng X, Wang X, Yang N, et al. (2021) A genetic variant conferred high expression of CAV2 promotes pancreatic cancer progression and associates with poor prognosis. Eur J Cancer 151: 94-105.
19. Liu Y, Wang M, Wang D, Fay WP, Korthuis RJ, et al. (2021) Elevated postischemic tissue injury and leukocyte-endothelial adhesive interactions in mice with global deficiency in caveolin-2: role of PAI-1. Am J Physiol Heart Circ Physiol 320(3): H1185-H1198.
20. Park H, Shin JA, Lim J, Lee S, Ahn JH, et al. (2022) Increased Caveolin-2 Expression in Brain Endothelial Cells Promotes Age-Related Neuroinflammation. Mol Cells 45(12): 950-962.
21. Jitariu AA, Raica M, Cimpean AM, Suciu SC (2018) The role of PDGF-B/PDGFR-BETA axis in the normal development and carcinogenesis of the breast. Crit Rev Oncol Hematol 131: 46-52.

22. Zhang Y, Cedervall J, Hamidi A, Herre M, Viitaniemi K, et al. (2020) Platelet-Specific PDGFB Ablation Impairs Tumor Vessel Integrity and Promotes Metastasis. Cancer Res 80(16): 3345-3358.

23. Jang M, Hara S, Kim GH, Kim SM, Son S, et al. (2021) Dutomycin Induces Autophagy and Apoptosis by Targeting the Serine Protease Inhibitor SERPINB6. ACS Chem Biol 16(2): 360-370.

ISSN: 2574-1241

DOI: 10.26717/BJSTR.2026.66.010290

Xingjun Guo . Biomed J Sci & Tech Res

Submission Link: https://biomedres.us/submit-manuscript.php