BMC Cancer

RESEARCH

Open Access

Machine learning based identification of an amino acid metabolism related signature for predicting prognosis and immune microenvironment in pancreatic cancer

Xiaohong Liu[1,2,3†], Xing Wang[1,2,3†], Jie Ren[1,2,3†], Yuan Fang[1,2,3], Minzhi Gu[1,2,3], Feihan Zhou[1,2,3], Ruiling Xiao[1,2,3], Xiyuan Luo[1,2,3], Jialu Bai[1,2,3], Decheng Jiang[1,2,3], Yuemeng Tang[1,2,3], Bo Ren[1,2,3*], Lei You[1,2,3*] and Yupei Zhao[1,2,3*]

Abstract

Background Pancreatic cancer is a highly aggressive neoplasm characterized by poor diagnosis. Amino acids play a prominent role in the occurrence and progression of pancreatic cancer as essential building blocks for protein synthesis and key regulators of cellular metabolism. Understanding the interplay between pancreatic cancer and amino acid metabolism offers potential avenues for improving patient clinical outcomes.

Methods A comprehensive analysis integrating 10 machine learning algorithms was executed to pinpoint amino acid metabolic signature. The signature was validated across both internal and external cohorts. Subsequent GSEA was employed to unveil the enriched gene sets and signaling pathways within high- and low-risk subgroups. TMB and drug sensitivity analyses were carried out via Maftools and oncoPredict R packages. CIBERSORT and ssGSEA were harnessed to delve into the immune landscape disparities. Single-cell transcriptomics, qPCR, and Immunohistochemistry were performed to corroborate the expression levels and prognostic significance of this signature.

Results A four gene based amino acid metabolic signature with superior prognostic capabilities was identified by the combination of 10 machine learning methods. It showed that the novel prognostic model could effectively distinguish patients into high- and low-risk groups in both internal and external cohorts. Notably, the risk score from this novel signature showed significant correlations with TMB, drug resistance, as well as a heightened likelihood of immune evasion and suboptimal responses to immunotherapeutic interventions.

Conclusion Our findings suggested that amino acid metabolism-related signature was closely related to the development, prognosis and immune microenvironment of pancreatic cancer.

Keywords Amino acid metabolism, Prognosis, Immune microenvironment, Pancreatic cancer, Machine-learning

†Xiaohong Liu, Xing Wang and Jie Ren contributed equally to this work.

*Correspondence:
Bo Ren
berserker94@163.com
Lei You
florayo@163.com
Yupei Zhao
zhao8028@263.net
Full list of author information is available at the end of the article

Introduction

Pancreatic cancer is an extremely lethal malignancy with an unfavorable 5-year survival rate of less than 10% [1]. It is frequently diagnosed in advanced stages, leading to restricted treatment choices and grim prognosis [2]. The aggressive nature of pancreatic cancer is attributed to its rapid progression, early metastasis, and resistance to conventional chemical therapies [3, 4]. Tumor microenvironment in pancreatic cancer is distinguished by dense fibrosis, immune suppression, and limited vascularization, which contribute to the aggressive behavior of the disease [5, 6]. Therefore, identifying reliable prognostic biomarkers and understanding the molecular mechanisms underlying pancreatic cancer progression and the interactions between tumor cells and the tumor microenvironment is essential for the development of effective therapeutic strategies.

Recent research has underscored the importance of amino acid metabolic alterations in the onset and advancement of cancer [7]. The primary role of amino acids is to serve as substrates for the synthesis of proteins. Amino acids can provide nitrogen and carbon skeleton for rapid proliferation of tumor cells, as well as metabolic intermediates for biosynthesis and energy production [8]. It can also support the needs of tumor progression by maintaining Redox balance and modulating epigenetic inheritance [9]. Furthermore, amino acid metabolism influences the immune microenvironment, which is conducive to tumor immune evasion and immunosuppression [10]. Alterations in amino acid metabolism have been shown to contribute to tumor growth, metastasis, and resistance to therapy [11]. Specifically, Glutamine provides nitrogen for nucleotide and amino acid biosynthesis, while maintaining mitochondrial membrane potential and redox balance [12]. Arginine is a precursor of nitric oxide biosynthesis, which promotes wound healing and the release of insulin-like growth factor 1, and exhibits multiple immunomodulatory effects [13]. As an important one-carbon donor, serine promotes tumor cell growth by participating in nucleotide synthesis [14]. Branched chain amino acids such as valine, leucine, and isoleucine can also supply nitrogen for the synthesis of important biological macromolecules like nucleotides [15]. Hence, targeting pathways related to amino acid metabolism has emerged as a promising therapeutic strategy for tumor treatment. Machine learning methods have been widely used in cancer research to identify prognostic biomarkers and predict patient outcomes [16, 17]. In the present study, we utilized an integration of machine learning algorithms to screen amino acid metabolism-related genes and construct a prognostic model for pancreatic cancer. The association between the prognostic model and the immune microenvironment was also explored to gain insights into the potential mechanisms underlying the role of amino acid metabolism in pancreatic cancer.

Given the dismal prognosis and limited treatment options for pancreatic cancer, there is a critical necessity to identify novel prognostic biomarkers and therapeutic targets. Recently, there has been growing interest in exploring the role of amino acid metabolism in cancer, including pancreatic cancer, as a potential target for therapy and prognostic prediction. This study endeavors to enhance the comprehension of pancreatic cancer through revealing the association of amino acid metabolism-related genes with patient prognosis and the immune microenvironment using machine learning techniques.

Method and materials

Data collection

Gene expression data from RNA-seq and relevant clinical information were obtained from TCGA (https://portal.gdc.cancer.gov/), ICGC (https://dcc.icgc.org/), and GTEx database (https://commonfund.nih.gov/GTEx). Microarray gene expression data and clinical information from studies such as GSE28735, GSE57495, GSE62452, and GSE85916 were collected from the GEO database (https://www.ncbi.nlm.nih.gov/geo/). Single cell transcriptomic dataset GSE205013 with 6 primary PDAC tumor samples and 9 liver metastatic tumor samples was also downloaded from GEO database. In this study, a cohort consisting of 690 patients with both clinical and gene expression data sourced from TCGA, ICGC, GSE57495, GSE28735, GSE85916 and GSE62452 was aggregated. Subsequently, batch effects were mitigated by normalizeBetweenArrays and removeBatchEffect function, and the cohort was randomly partitioned into a training set (n = 487) and a validation set (n = 203) to address the potential bias inherent in individual datasets and avoid overfitting, which subsequently subjected to the integration of machine learning algorithms framework for the identification of prognostic genes.

Identification of amino acid metabolic signature

The ssGSEA (single-sample Gene Set Enrichment Analysis) and WGCNA (Weighted Gene Co-expression Network Analysis) algorithms were utilized to screen out the co-expressed genes of amino acid metabolism and pancreatic cancer ($p < 0.05$). The ScWGCNA analysis was also conducted to identify gene modules associated with amino acid metabolism as well as pancreatic cancer in single cell transcriptomic data through hdWGCNA R package. hdWGCNA [18] is a framework tool designed for analyzing co-expression networks in high-dimensional transcriptomic data, providing capabilities for network inference, module identification, and

data visualization. The differentially expression analysis was conducted by limma R package with |logFC|>1 & adj.$p<0.05$ in cohort comprising of patients from TCGA and GTEx datasets to balance the need for biological significance with the risk of overlooking potentially meaningful research subjects. Then, 40 sharing genes were obtained by intersection of amino acid metabolism related genes, pancreatic cancer co-expressed genes and differentially expressed genes via veen plot. Univariate cox regression analysis was utilized to identify potential prognostic biomarkers in the investigation. Finally, 13 genes with significant prognostic potential were put into the framework of an integration of machine learning algorithms to further screen out the genes with leading AUC value and C-index. Machine learning (ML) empowers systems to derive insights from data, recognize patterns, and execute decisions with limited human oversight. The fundamental tenet of machine learning is to facilitate the training of models utilizing datasets. This iterative learning process enables the model to extract and internalize the intrinsic patterns and interrelationships present within the data. The machine learning method in this study is an integration and combination with varying hyperparameters of ten machine learning algorithms including CoxBoost, Lasso, stepwise Cox, plsRcox, Ridge, Enet, SurvivalSVMS, GBMs, SuperPC and RSF [19].

Construction and validation of a novel prognostic model

Utilizing machine learning algorithms and multi-Cox regression analysis, we identified a four-gene signature associated with amino acid metabolism that demonstrated significant prognostic value. Subsequently, the coefficients for each gene were computed. The amalgamated dataset from TCGA and ICGC was designated as the training set, and the risk score was calculated using the formula: $Risk\ score = \Sigma\ (Coefi \times Exp)$. A prognostic model was then established, and internal validation was conducted using TCGA and ICGC as separate validation sets. For external validation, GSE85916 was employed as a validation set with a limited sample size, while a combined cohort from GSE85916, GSE28735, GSE62452, and GSE57495 was utilized as an external dataset with a larger sample size. The new signature was verified by applying the same calculation formula and cut-off point.

GSEA analysis of the amino acid metabolism related signature

GSEA analyses was performed via clusterProfiler, DOSE and enrichplot R packages to reveal the GO, KEGG pathways and Hallmarks between high- and low- risk groups. Then, the correlation between the expression levels of 4 prognostic signatures and cancer hallmarks was calculated through ggcor R package. The KEGG pathways altered between the high- and low-expression groups of the four genes were analyzed via LinkedOmics (https://www.linkedomics.org/login.php).

Correlation analyses of risk scores with TMB, drug sensitivity and tumor immune microenvironment

Single Nucleotide Polymorphism information was downloaded from TCGA database. Maftools, ggpubr and reshape2 R packages were used to analyze the alteration of tumor mutation burden in high- and low- risk groups. Then, the alteration of drug sensitivity in high- and low-risk groups was investigated by oncoPredict R package. CIBERSORT algorithm and ssGSEA analysis were applied to reveal the association of risk scores with tumor immune infiltration, immune escape and immune therapy. Estimate R package was also applied to reveal the changes between tumor microenvironment and risk scores.

Single-cell and bulk RNA_seq data analysis of the differentially expression and prognostic value of 4 selected signatures

The TISCH database (http://tisch1.comp-genomics.org/) was employed to investigate the expression of four genes related to amino acid metabolism in various tumor cell populations and to assess the association between their expression levels and prognosis using single-cell transcriptome data. Additionally, bulk RNA data from the TCGA was analyzed to confirm the expression alterations and prognostic significance of the identified signatures.

Immunohistochemistry, qPCR and transwell analyses

Tissue microarrays consisting of tumor tissues and paired normal tissues were collected from 40 patients who underwent curative resection for PDAC at Peking Union Medical College Hospital. The assessment of all tissue samples was carried out with the assistance of the ImageJ-based IHC Profiler plugin, followed by manual corrections. RNA was isolated from HPNE, PANC1, T3M4, and CFPAC1 cell lines employing TRIzol reagent (Takara, Japan). The immunohistochemistry assay and Real-time quantitative PCR analysis were conducted as previously reported [20]. The primer sequences can be found in Table 1. A suspension of 6.0–8.0×10^4 BxPC-3 cells was prepared in the corresponding medium devoid of fetal bovine serum (FBS) and introduced into the upper chamber of a Costar 3422 transwell system. For the invasion assay, the upper chamber was pre-coated with a 1:20 dilution of Matrigel; however, for the migration assay, the upper chamber remained uncoated. The lower chamber was filled with complete medium supplemented with 10% FBS. After 24 h, the migrated cells were

Table 1 Primer sequences in qPCR

Name	Forward primer (5′−3′)	Reverse primer (5′−3′)
GAPDH	ACCCACTCCTCCACCTTT	CTGTTGCTGTAGCCAAATTCGT
GFPT2	AGGTGCATTCGCGCTGGTT	TGTGGAGAGCTTGTATTTGCTCC
P4HA1	AATCTGGTCCTTAAGGATATG TCAG	GGAAGATTACCCTTTGAG ATGGT
HAAO	CAGGCAGGCACACCACTCAG	CCCATTGTCACCACCGAGGAG
SAT2	AGGTTGCCCTGAATAAGGGTT	CTCGCCTTCAAATCGAAAGGA

fixed with methanol for 30 min and subsequently stained with a 0.5% crystal violet solution for 30 min. Following staining, the cells were rinsed, dried, and the number of stained cells in each filter was quantified under a microscope. Each experimental group was analyzed using three replicates.

Statistical analysis

Kaplan–Meier analysis and log-rank tests were utilized to compare the overall survival rates between high- and low-risk cohorts. Furthermore, univariate and multivariate Cox regression analyses were executed to ascertain independent prognostic factors influencing survival duration. The Wilcoxon test was employed to discern variations in gene expression, TMB, drug responsiveness, immune scores, IPS, and TIDE scores across the two groups. Spearman correlation analysis was performed to evaluate the interrelation between these cohorts. Image J was used for data analysis of IHC and migration and invasion experiments. All statistical assessments were carried out using R software (version 4.3.0). A two-tailed p-value of $p < 0.05$ was considered statistically significant, denoted by "***" for $p < 0.001$, "**" for $p < 0.01$, and "*" for $p < 0.05$.

Results

Integrative analysis of bulk and single cell transcriptomic data identified amino acid metabolism related signature in pancreatic cancer

Amino acids can be used as an energy source and are involved in biosynthesis. Amino acid derivatives are crucial in epigenetic regulation and immune responses linked to tumor initiation and metastasis. In order to screen out amino acid metabolism related genes, both bulk and single cell transcriptomic data were analyzed.

The ScRNA_Seq data included 6 primary PDAC tumor samples and 9 liver metastatic tumor samples. By combining automated annotation with single R and manual annotation, cells from GSE205013 were annotated into 7 cell subtypes, including Epithelial cells, NK cells, Macrophages, fibroblasts, endothelial cells, B cells and T cells (Fig. 1A-C). These cells were subjected to hdWGCNA analysis, focusing on the characteristic genes of Epithelial cells for module clustering. With the soft thresholding power β established at 12, the fit index for scale-free topology achieved 0.8, leading to the identification of 18 modules (Fig. 1D-F). Subsequently, the correlation between identified modules and traits such as pathways in cancer, pancreatic cancer, cell cycle, and amino acid metabolism pathways was explored (Fig. 1G). Modules Epi-M7, Epi-M15, Epi-M16, Epi-M17, and Epi-M18, showing significant positive correlations with pathways in cancer, pancreatic cancer, cell cycle, and close associations with amino acid metabolism, were individually extracted. Differential expression analysis was conducted on these extracted modules in primary and liver metastasis samples to identify genes related to liver metastasis and amino acid metabolism (Fig. 2D). The ssGSEA and WGCNA algorithms were also applied to estimate amino acid metabolism-related signaling pathways and screen related co-expression gene modules in bulk RNA_Seq data. The results revealed that MElightcyan and MEgrey were significantly positively correlated with majority of amino acid metabolic pathways (Fig. 2A). Then, genes co-expressed with pancreatic cancer and the differentially expressed genes in PAAD were also identified (Fig. 2B, C). Intersection of 1049 genes related to amino acid metabolism, 3887 DEGs, and 6787 co-expressed genes in PAAD were conducted. Forty shared genes underwent the subsequent analysis (Fig. 2E). Unicox regression analysis further identified 13 prognostic genes for pancreatic cancer (Fig. 2F). The AUC values of 89 kinds of machine learning combination and the C-index values of 99 kinds of machine learning methods were further calculated by LOOCV framework with varying hyperparameters of ten machine learning algorithms including CoxBoost, Lasso, stepwise Cox, plsRcox, Ridge, Enet, SurvivalSVMS, GBMs, SuperPC and

(See figure on next page.)

Fig. 1 Identification of amino acid metabolism related signature based on single cell transcriptomic data. **A** Cell score heatmap showed cell subgroups annotated by Single R package. **B** Bubble map showed manual annotation and correction of each cell subpopulation in different clusters. **C** UMAP and tSNE dimensionality reduction and clustering showed 7 different cell types. **D** This section describes related parameters of WGCNA analysis. **E** Dendrogram of different modules. **F** Display of the 18 hdWGCNA modules. **G** Correlation of different modules with signaling pathway traits

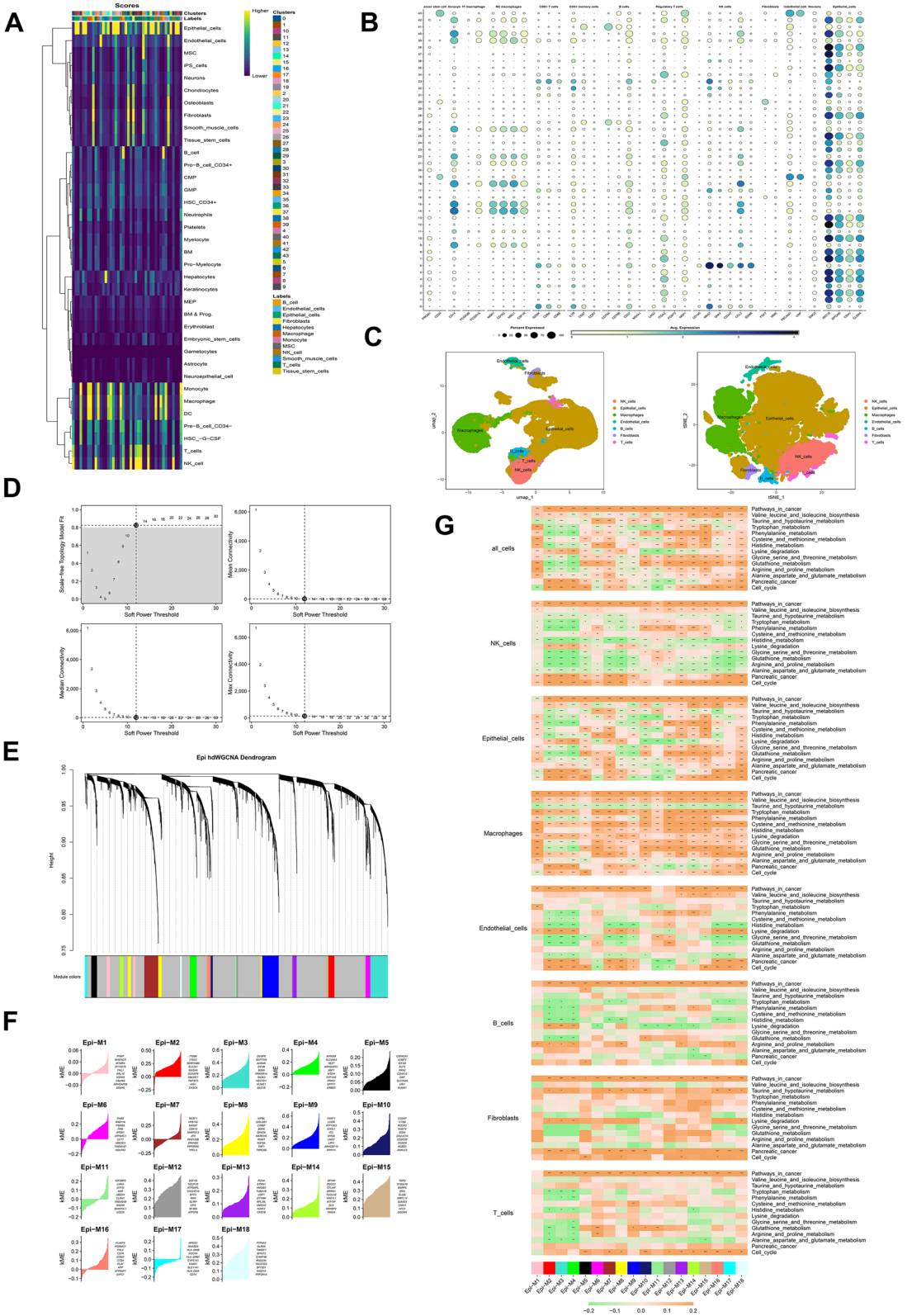

Fig. 1 (See legend on previous page.)

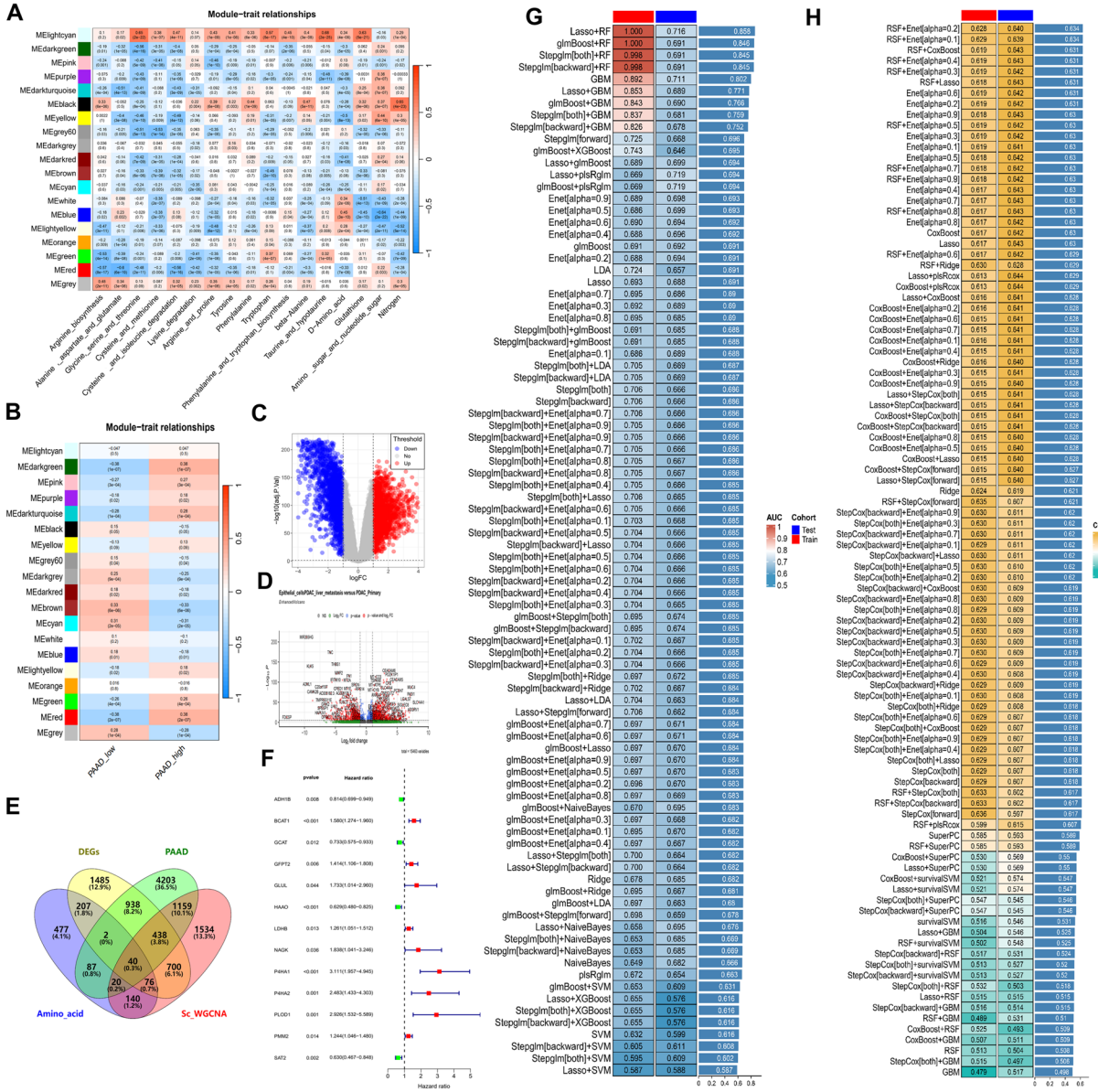

Fig. 2 Identification of amino acid metabolism related signature based on bulk RNA_Seq data. **A** WGCNA analysis identified gene modules corelated with amino acid metabolic signal pathways. **B** WGCNA analysis of co-expressed genes with pancreatic cancer. **C** Volcano map showed differentially expressed genes between the tumor group and the control group in the TCGA-PAAD dataset. **D** Enhanced Volcano map displayed the differentially expressed genes between primary and liver metastatic groups in ScRNA_Seq data. **E** Venn diagram revealed 40 potential prognostic genes in pancreatic cancer. **F** Forest map showed that 13 genes had significant prognostic value. **G** An integration of 89 machine learning analyses screened prognostic genes with the greatest AUC values. **H** The utilization of 99 machine learning algorithms to screen for prognostic genes with the highest C-index values

RSF (Fig. 2G, H). The results suggested that when the integrative model was Lasso + RF, the AUC value was the highest (0.858), and when the model combination was RSF + Enet[alpha = 0.2] or RSF + Enet[alpha = 0.1], it achieved the highest C-index (0.634).

Construction and validation of a novel prognostic model

Four amino acid metabolism related genes including GFPT2, P4HA1, HAAO and SAT2 were further screened out based on the optimal models with leading AUC and C-index to develop a novel prognostic signature for pancreatic cancer. The coefficients of the 4 genes were calculated via multi-cox regression analysis. The formula

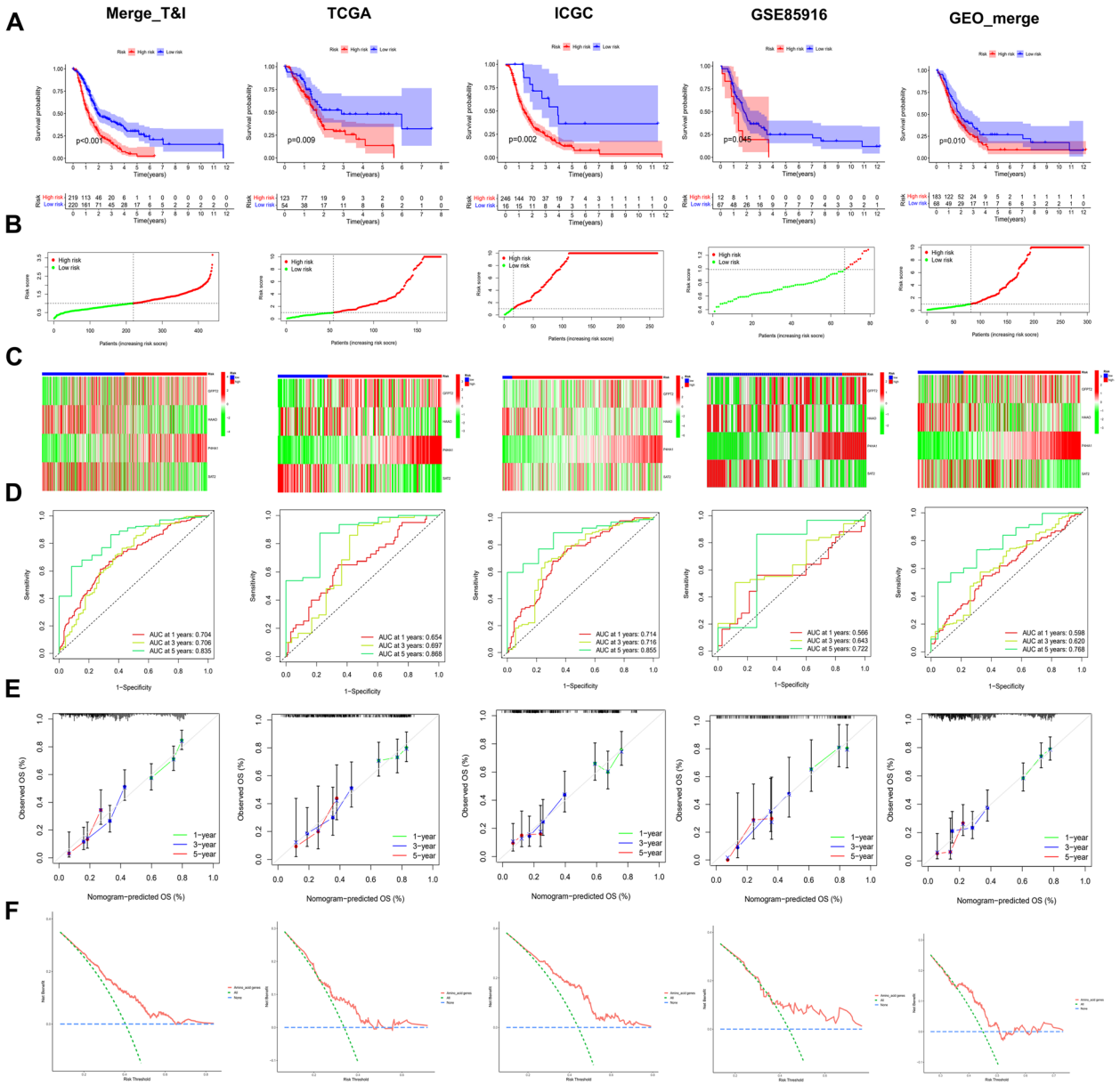

Fig. 3 Construction and validation of an amino acid metabolic prognostic model. **A** A prognostic model of 4 amino acid metabolic genes was constructed in a merged dataset comprising of TCGA and ICGC cohorts, which was validated in the internal cohorts of ICGC and TCGA, as well as the external cohorts of GSE85916 and the combined GEO dataset. **B** Patients were divided into high and low risk groups based on risk scores calculated by the prognostic model. **C** Heatmaps showed the expression of four prognostic risk genes in the high-low risk groups. **D** ROC curves, **E** Calibration curves and **F** DCA curves further evaluated the accuracy, robustness and clinical value of the prognostic model

of *Risk score* = *0.4352* × *GFPT2 Exp-0.2725* × *HAAO Exp* + *1.3791* × *P4HA1 Exp—0.8734* × *SAT2 Exp* was used to calculate the risk score of each patient, thus dividing patients into high and low risk groups in the train dataset, which was validated in different internal and external data sets (Fig. 3A-C). The overall survival time of patients with higher risk scores was significantly shorter than that in the low- risk group (Fig. 3A). Subsequently, the risk scores of the training set and each validation set showed superior discriminative ability for patient outcomes at 1, 3, and 5 years, as indicated by the ROC curves (Fig. 3D). Additionally, the calibration curves for different datasets at 1, 3, and 5 years demonstrated the robustness of the prognostic model, further confirming its high reliability (Fig. 3E). Furthermore, by comparing with the C-Index of six previously published prognostic models related to

pancreatic cancer, we found that the amino acid-related signature exhibited the highest C-Index in both the training and validation sets. This suggests that our signature demonstrates superior robustness and consistency in the prognostic prediction of pancreatic cancer (Supplementary Fig. 1). The DCA curve of the amino acid metabolism-related signature to predict the prognosis of pancreatic cancer showed that when the survival threshold of the training set and the internal validation set exceeded 0.2, and the survival threshold of the external validation set exceeded 0.4, this feature was beneficial to predicting patients' prognosis (Fig. 3F).

GSEA analysis was performed in high- and low- risk groups

To further explore the gene sets and signaling pathways enriched in each risk groups, GSEA analysis was conducted. The Gene Ontology analysis revealed significant enrichment of biological processes such as cornification, DNA geometric change, and keratinization in the high-risk cluster, while processes like digestion, peptide hormone secretion, and regulation of hormone levels were significantly promoted in the low-risk group (Fig. 4A, B & Supplementary Table 2). The KEGG pathway enrichment analysis showed that signaling pathways including adherens junction, ECM receptor interaction, focal adhesion, and pancreatic cancer were significantly active in the high-risk cluster, whereas pathways including chemokine signaling and drug metabolism cytochrome P450 were notably decreased (Fig. 4C, D & Supplementary Table 3). Furthermore, the Hallmarks pathway enrichment analysis indicated significant enrichment of signaling pathways like the HIF1 signaling pathway, ECM receptor interaction, and pancreatic cancer in the high-risk group, while pathways like steroid hormone biosynthesis and insulin secretion were obviously elevated in the low-risk group (Fig. 4E, F & Supplementary Table 4). These findings suggest that relevant signaling pathways associated with pancreatic cancer are active in the high-risk cluster, further validating the predicting accuracy of the signature in pancreatic cancer.

The correlation analyses of risk scores with TMB and drug sensitivity

TMB analysis in high- and low-risk subgroups of patients revealed a mutation rate of 85% in the high-risk cluster, compared to 64.29% in the low-risk cluster (Fig. 5A, B). The three most common types of mutations in both groups were missense mutation, nonsense mutation, and frame shift del (Fig. 5A, B). Notably, there is a significant positive association between risk scores and TMB ($R=0.22$, $p=0.0091$), and TMB was significantly upregulated in the high-risk group compared to the low-risk group (Fig. 5C, D). Combining TMB and risk scores in survival analysis shows that patients with low TMB and low risk have a better prognosis than those with high TMB and high risk (Fig. 5E, F). In the high-risk group, the mutation rates of KRAS, CDKN2A, and TP53 showed a significant increase compared to the low-risk cluster, whereas the SMAD4 mutation load showed a significant decrease (Fig. 5G-J). Moreover, to investigate the relationship between risk scores and drug resistance in pancreatic cancer patients, we analyzed the IC50 values of prevalent chemotherapy and targeted drugs in high- and low-risk groups. The results reveal decreased sensitivity to drugs such as 5-Fluorouracil, Irinotecan, Gemcitabine, and KRAS(G12C) Inhibitor-12 in high-risk subgroup, implying a significant correlation between high-risk scores and drug resistance (Fig. 5K-R). Additionally, we utilized the GDSC2 website (https://www.cancerrxgene.org/) to predict the expression of four amino acid metabolism-related prognostic genes in relation to potential drug sensitivity. The results indicated that the high expression groups of GFPT2 and P4HA1 exhibited more pronounced drug resistance, whereas the high expression groups of HAAO and SAT2 demonstrated lower drug resistance. This suggests a detrimental prognostic role for GFPT2 and P4HA1(Supplementary Fig. 2).

Immune infiltration and immune function analyses in high- and low- risk groups

This study also delved into the impact of amino acid metabolism-related risk scores on immune infiltration and immune function. Our findings reveal substantial correlations between four amino acid metabolism-related genes and the majority of immune checkpoints. The identification of immune checkpoints has paved the way for novel approaches in tumor immunotherapy. Notably, the risk scores exhibit a significant positive correlation with the immune checkpoint CD274, CD80, and CD44 et al., while demonstrating a marked negative correlation with CTLA4, LAG3 and TNFRSF14, implying a potential link between the risk scores and tumor immune suppression (Fig. 6A). Furthermore, correlation analyses between immune infiltration levels and risk scores unveil a negative association with immune cells of positive immune function, including $CD8^+$ T cells, NKT cells, and $CD4^+$ Th1 cells, while demonstrating a positive correlation with inhibitory immune cells like cancer-related fibroblasts, $CD4^+$ Th2 cells, and Treg cells (Fig. 6B). Subsequent ssGSEA analysis of immune cells in pancreatic cancer further indicates a significant reduction in infiltrating levels of activated $CD8^+$ T cells, activated B cells, CD56 bright NK cells, CD56 dim NK cells, Th17 cells, effector memory $CD4^+$ T cells, and effector memory $CD8^+$ T cells in the high-risk subgroup, alongside a notable increase in infiltrating levels of activated $CD4^+$ T cells,

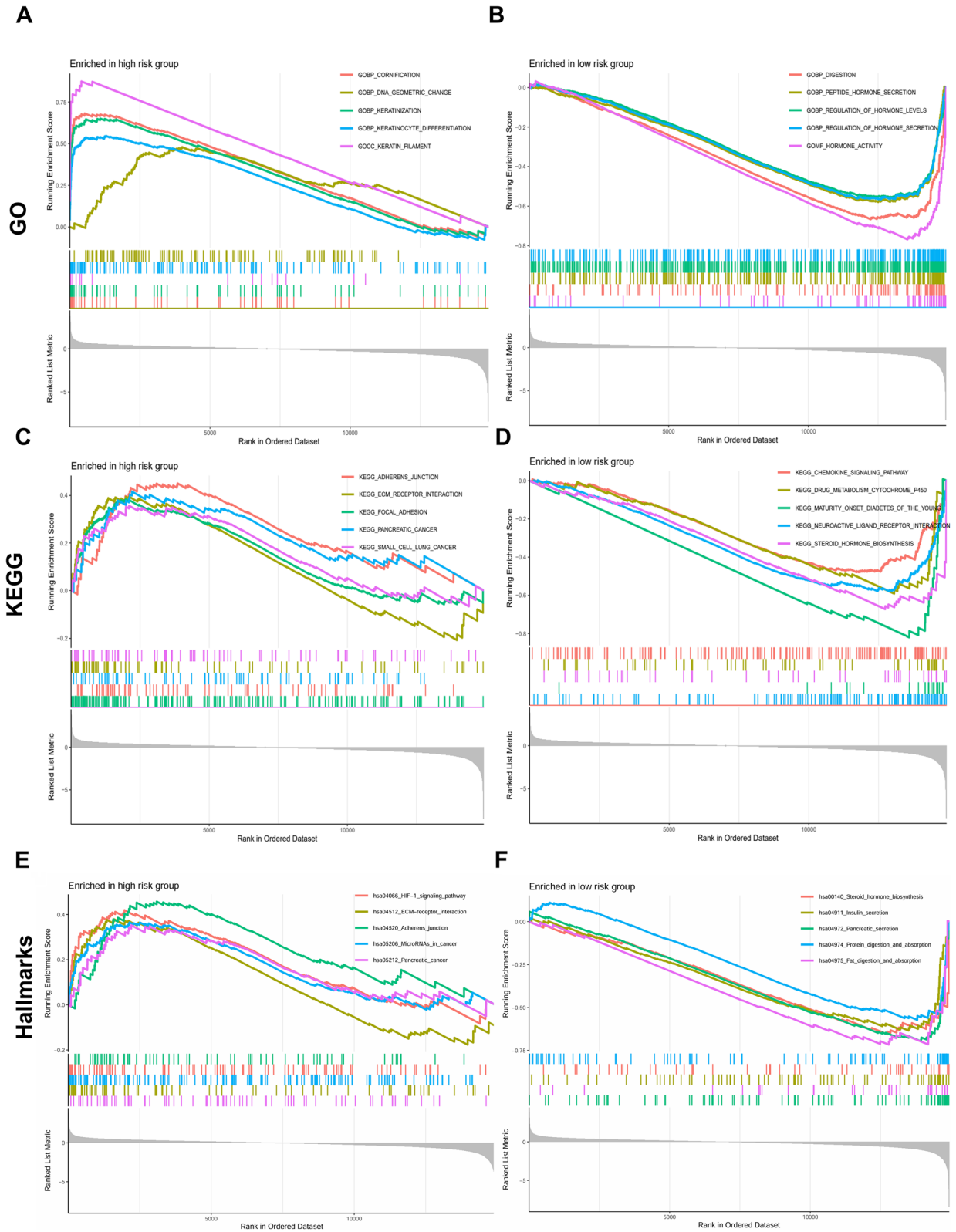

Fig. 4 GSEA analysis between high- and low- risk groups. **A**, **B** GO analysis, **C**, **D** KEGG analysis and **E**, **F** Hallmarks of high- and low-risk groups

activated dendritic cells, Th2 cells, and immature dendritic cells (Fig. 6C, D). Moreover, results from immune function-related analyses suggest a substantial positive correlation between risk scores and APC_co-inhibition, while demonstrating a significant negative correlation with T cell co-stimulation, cytolytic activity, and Type II IFN response (Fig. 6E, F). Taken together, these findings collectively indicate an association of amino acid-related risk scores with an immune-suppressive microenvironment.

Immune escape and immune therapy analyses in different groups

By further investigating the impact of amino acid metabolism-related risk scores on immune escape, immunotherapy, and the immune microenvironment in pancreatic cancer, it was found that the risk scores were significantly positively correlated with tumor microenvironment scores (Fig. 7A). Moreover, higher risk scores were related to higher PDL1 expression and PDL1 pathway in cancer (Fig. 7B, C). In the high-risk group, the IPS scores for CTLA4(-)PD1(+), CTLA4(+)PD1(+), CTLA4(+)PD1(-), and CTLA4(-)PD1(-) were significantly decreased compared to the low-risk cluster, indicating poor response to immunotherapy in high-risk patients (Fig. 7D-G). Furthermore, the TIDE scores, exclusion, MDSC, CAF, and CD274 were significantly higher in the high-risk group, while the scores of dysfunction and CD8 were significantly higher in the low-risk group. MSI showed no significant correlation with risk scores (Fig. 7I-P). Based on Fisher test analysis, there was a significant correlation between the risk scores and patient responsiveness to immune therapy. Further submap analysis suggested that patients with low-risk rather than high-risk may be responsive to CTLA4 immune therapy (Adjusted $p < 0.05$, Fig. 7Q, R). The results indicated the higher potential of immune escape and immune therapy resistance in high- risk group.

Correlation analyses of 4 prognostic genes with cancer hallmarks and the GSEA analysis

In-depth exploration was conducted into the correlation between the four amino acid metabolism-related signature and cancer-related hallmarks. The results revealed that the oncogenic GFPT2 and P4HA1 showed a significant positive correlation with most cancer hallmarks signaling pathways, while displaying a negatively significant association with fatty acid metabolism and oxidative phosphorylation (Fig. 8A, C). In contrast, the protective gene SAT2 exhibited a significant negative correlation with most cancer hallmarks, while showing a significant positive correlation with oxidative phosphorylation (Fig. 8B). HAAO showed a negative correlation with the MYC_targets signaling pathway, while revealing a significant positive association with the IL6-JAK-STAT, IL2-STAT5 signaling pathway, and angiogenesis (Fig. 8D). Moreover, ssGSEA analysis indicated that GFPT2 was significantly positively correlated with multiple immune-related signaling pathways, including the inflammatory response pathway, selective expression of chemokine receptors during T cell polarization, and development of pulmonary dendritic cells and macrophage subsets (Fig. 8E). P4HA1 was significantly positively correlated with biosynthesis of amino acids, HIF-1 signaling pathway, and pancreatic cancer (Fig. 8G). HAAO was positively correlated with Th1 and Th2 cell differentiation, Th17 cell differentiation, while negatively correlated with the cell cycle (Fig. 8F). SAT2 was significantly positively correlated with glycine, serine, and threonine metabolism, tryptophan metabolism, while showing a notably negative correlation with pancreatic cancer and proteoglycans in cancer (Fig. 8H). These findings offered insights into the potential mechanisms underlying the association between amino acid metabolic signature and cancer-related signaling pathways (Fig. 8 & Supplementary Fig. 6).

Validation of the four-gene amino acid metabolic signature

By analyzing the single-cell sequencing data from the TISCH database, it revealed the expression patterns and prognostic significance of GFPT2, P4HA1, HAAO, and SAT2 in different cell types within pancreatic cancer (Fig. 9A-J). GFPT2 was predominantly expressed in fibroblasts, while P4HA1 showed primary expression in DCs, macrophages, mast cells, fibroblasts, and malignant cells, with partial expression in other cell types (Fig. 9K, L). HAAO exhibited main expression in DCs, macrophages, mast cells, and fibroblasts, and SAT2 demonstrated

(See figure on next page.)

Fig. 5 Risk score was correlated with tumor mutation burden and drug sensitivity. **A**, **B** The top20 mutated genes in high- and low- risk groups. **C** The tumor mutation burden in high-risk group was significantly higher than in low-risk group. **D** Risk scores were significantly correlated with tumor mutation burden. **E**, **F** Tumor mutation burden and risk score were significantly correlated to unfavorable prognosis. **G** KRAS, (H)TP53, **I** CDKN2A and **J** SMAD4, as the four genes with the highest mutation rates in pancreatic cancer, showed significant changes in the high-low- risk groups. The drug sensitivity of (**K**) 5-Fluorouracil, **L** Irinotecan, **M** Sorafenib, **N** Gemcitabine, **O** Epirubicin, **P** KRAS(G12C) Inhibitor-12, **Q** Oxaliplatin and **R** Erlotinib were significantly declined in high-risk group than in low-risk group

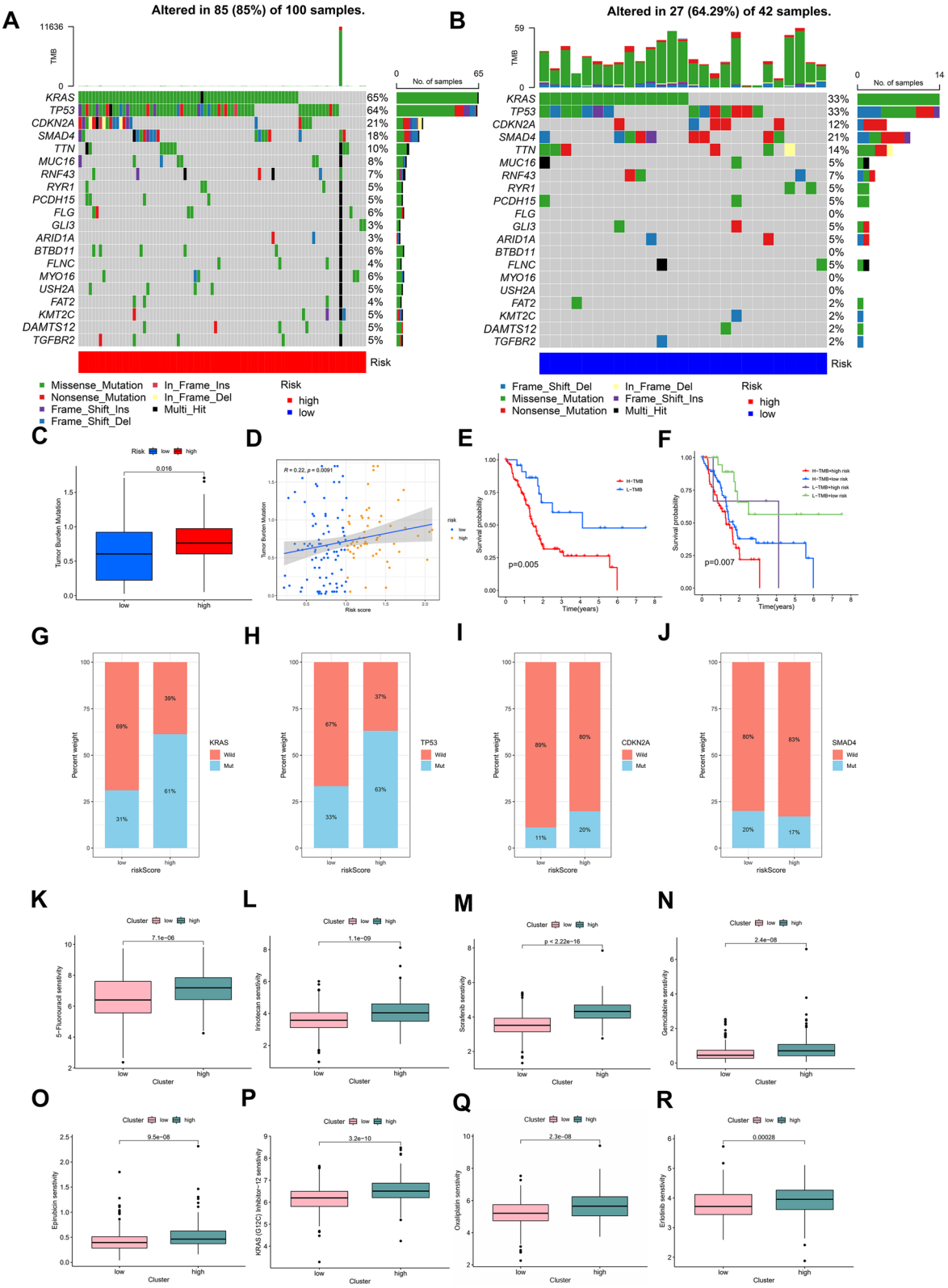

Fig. 5 (See legend on previous page.)

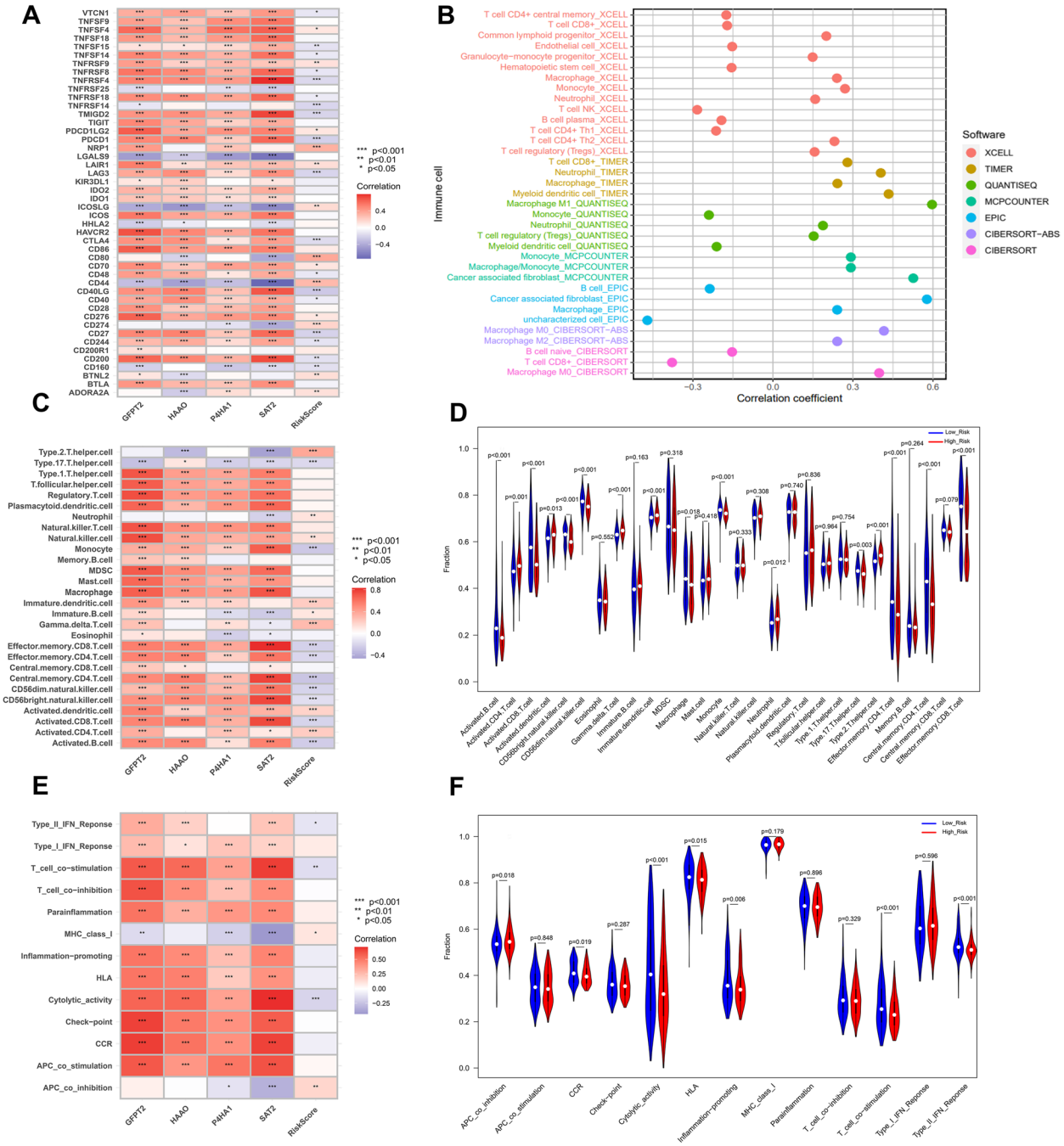

Fig. 6 The correlation of immune infiltration and immune function with risk scores. **A** The correlation of 4 prognostic genes and risk score with immune check points. **B** Immune infiltration analysis of risk scores from 7 software, including XCELL, TIMER, QUANTISEQ, MCPCOUNTER, EPIC, CIBERSORT-ABS and CIBERSORT. **C**, **D** The correlation analysis and GSEA analysis of immune infiltration in high- and low- risk groups. **E**, **F** The correlation analysis and GSEA analysis of immune function in high- and low-risk groups

expression in the majority of cell types (Fig. 9M, N). Furthermore, GFPT2 and P4HA1 were associated with an adverse prognosis, as indicated by a hazard ratio (HR) greater than 1 (Fig. 9O, P). Conversely, SAT2 displayed a protective prognostic effect with an HR less than 1 (Fig. 9Q). However, the prognostic value of HAAO was not significant within the single-cell sequencing dataset (Fig. 9L). These findings shed light into the expression profiles and prognostic implications of these genes across different cell populations in pancreatic cancer.

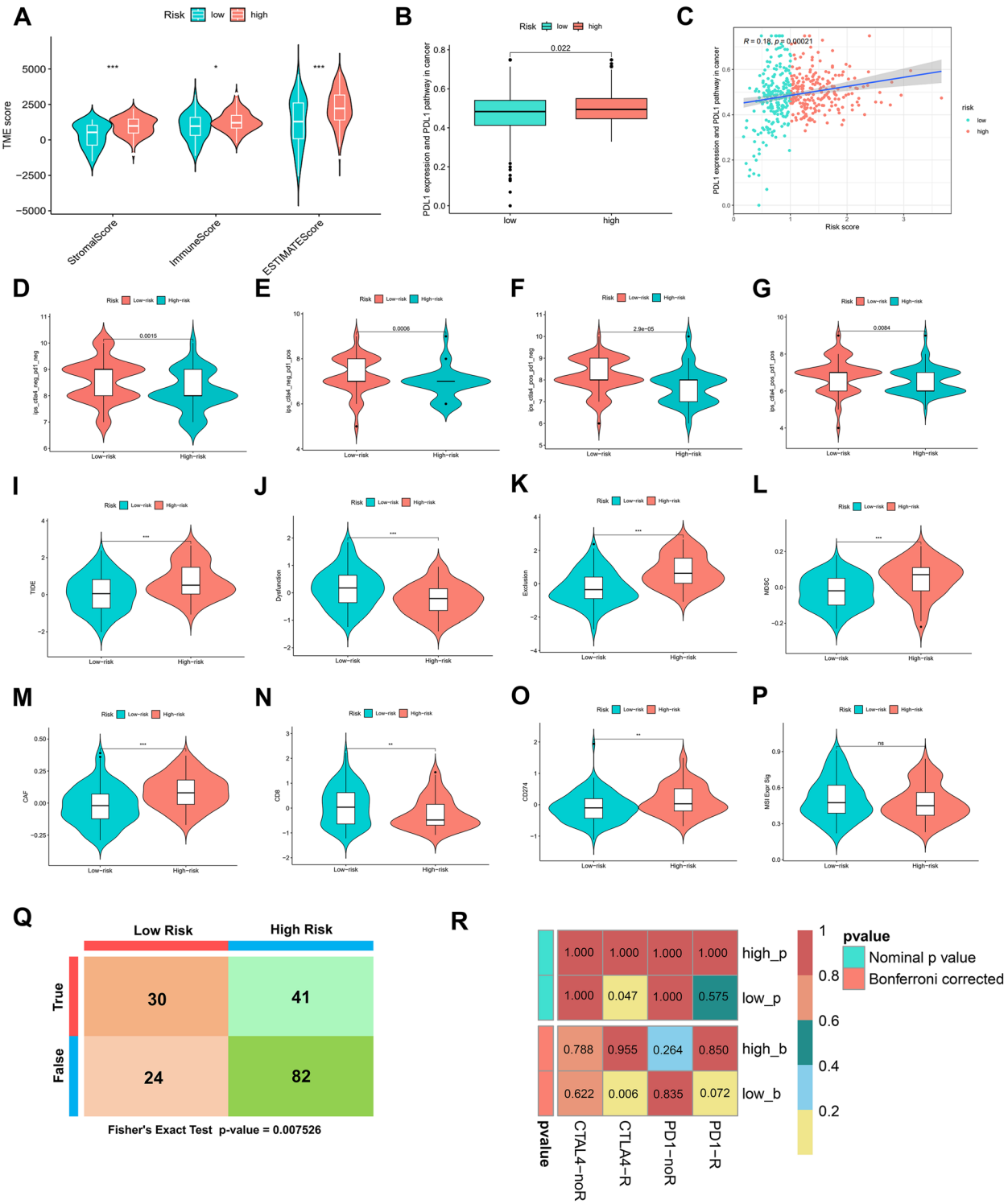

Fig. 7 The correlation of risk score with immune escape and immune therapy. **A** Stromal score, immune score and TME score in high- and low-risk groups. **B**, **C** The score of PDL1 expression and PDL1 pathway in cancer was significantly higher in high-risk group than in low-risk group. The IPS scores of (**D**) CTLA4(-)PD1(+), **E** CTLA4(-)PD1(-), **F** CTLA4(+)PD1(-) and (**G**) CTLA4(+)PD1(-) were significantly decreased in high-risk group. **I**-**O** The risk score was significantly associated with immune escape in PAAD. **P** MSI scores had no significant correlation in high- and low-risk groups. **Q** Fisher's test of immune therapy response in high- and low-risk groups. **R** Submap analysis identified significant response of CTLA4 in low-risk group

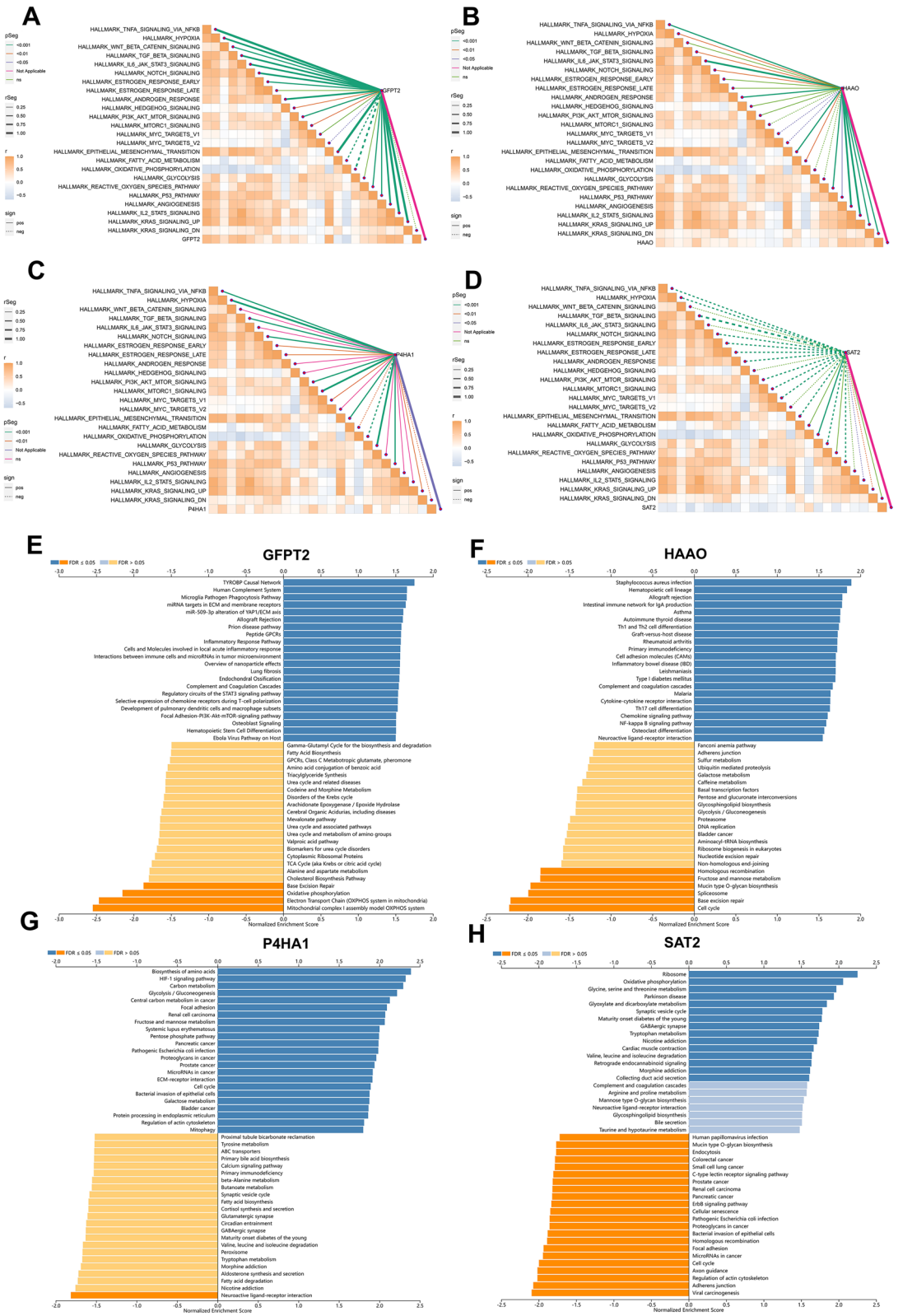

Fig. 8 The correlation of 4 prognostic genes with cancer hallmarks. Correlation analysis of (**A**) GFPT2, **B** HAAO, **C** P4HA1, and **D** SAT2 with cancer hallmark signaling pathways based on ggcor R package. GSEA analysis of (**E**) GFPT2, **F** HAAO, **G** P4HA1, and **H** SAT2 from Metascape database

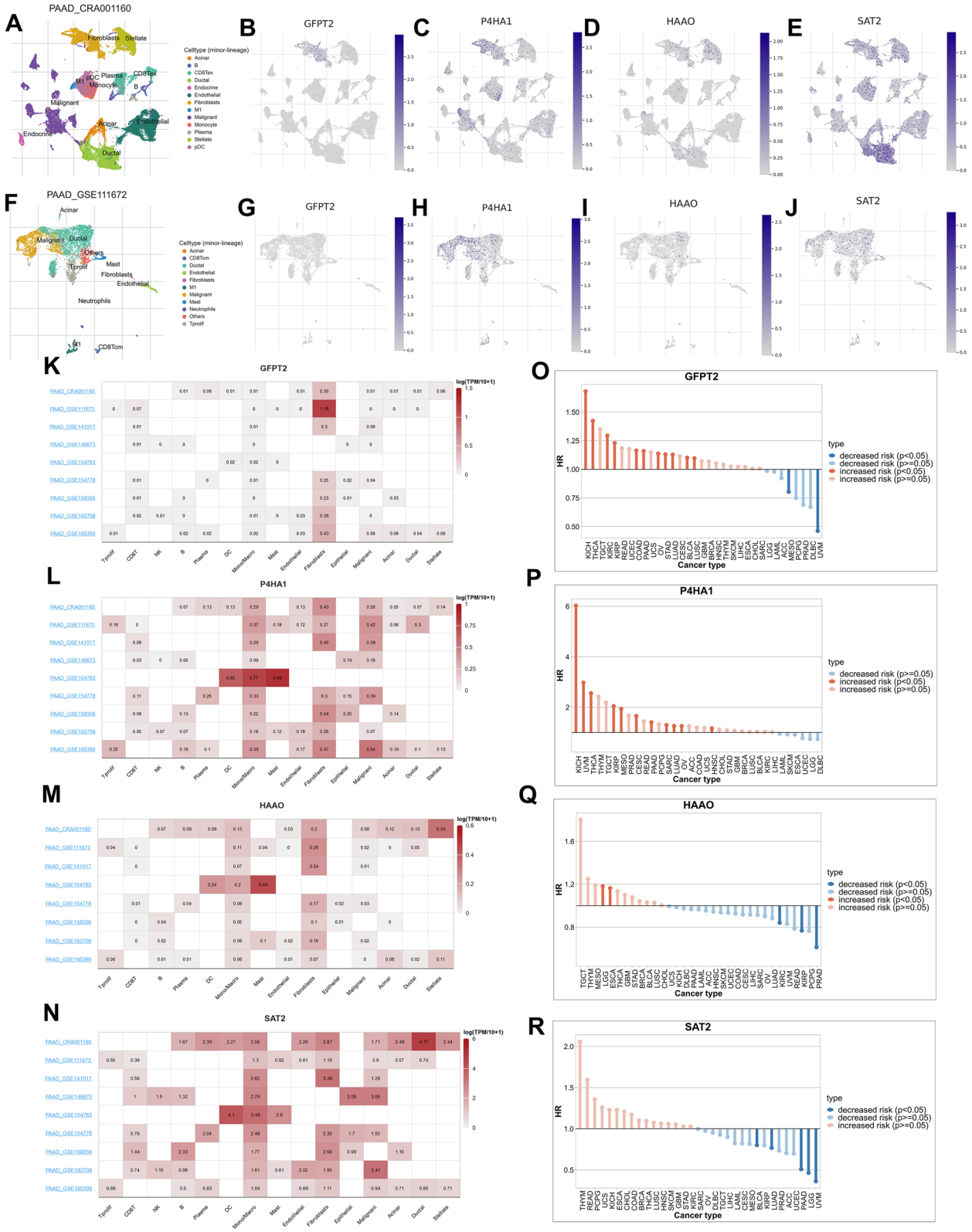

Fig. 9 Single-cell RNA-seq analysis identified the expression and prognostic value of 4 amino acid metabolic genes. **A** Minor-lineage annotation of cell subgroups in PAAD-CRA001160. **B** GFPT2, **C** P4HA1, **D** HAAO, and **E** SAT2 expression in PAAD-CRA001160. **F** Minor-lineage annotation of cell subgroups in PAAD-GSE11672. **G** GFPT2, **H** P4HA1, **I** HAAO, and **J** SAT2 expression in PAAD-GSE11672. **K**, **L**, **M**, **N** The expression levels of GFPT2, P4HA1, HAAO, and SAT2 in each cell type. **O**, **P**, **Q**, **R** The prognostic value of GFPT2, P4HA1, HAAO, and SAT2 in each cancer

Bulk RNA-seq data from the TCGA was also investigated to validate the differential expression and prognostic ability of the four amino acid metabolism-related genes. The results revealed that high expression of GFPT2 and P4HA1 was related to adverse prognosis, while elevated expression of HAAO and SAT2 indicated longer survival time (Fig. 10A-H). These findings align with the results of our prognostic model analysis. Further Cox regression analysis for the four amino acid metabolism-related genes and a comparison of the survival curves between high- and low-risk groups also demonstrated a fundamental consistency with our initial results (Supplementary Fig. 3–4). Furthermore, qPCR was applied to assess the expression disparities of four promising prognostic targets in the human normal pancreatic cell line HPNE and human pancreatic cancer cell lines PANC1, T3M4, and CFPAC1. The findings revealed a significant up-regulation of GFPT2 and P4HA1 RNA expressions in pancreatic cancer cell lines, while the converse was observed for HAAO and SAT2. These results align with our analytical expectations. (Fig. 10I-L). Using immunohistochemistry, the protein expression of four potential targets was examined between PDAC (n = 40) and normal tissues (n = 40). Subsequently, based on their expression levels, the two groups were categorized as negative (-), low positive (+), positive (++), and high positive (+++). All four amino acid-related targets were expressed in the cytoplasm and cell membrane. GFPT2 exhibited positive rate of 92.5% (37/40) in the PDAC and 5% (2/40) in the normal group. P4HA1 showed positive rate of 100% (40/40) in both PDAC and normal groups, with the low positive rate in the normal group being 80% (32/40) and the positive rate in PDAC group being 90% (36/40). HAAO displayed positive rate of 100% (40/40) in the normal group and 65% (26/40) in PDAC. SAT2 exhibited positive rate of 100% (40/40) in the normal group and 85% (34/40) in PDAC. The Wilcoxon test results indicated that the positive areas of GFPT2 and P4HA1 in PDAC were notably elevated than those in the normal group, while reversed in HAAO and SAT2 ($p < 0.001$, Fig. 10M, N). Additionally, the correlation between the positive area and survival time was examined in 22 out of 40 patients with integrate follow-up information. The results revealed a negative association between GFPT2 and P4HA1 expression and progress free survival (PFS) time in PDAC patients, whereas the expression levels of HAAO and SAT2 were positively associated with PFS time in PAAD patients. (Adjusted $p < 0.05$, Fig. 10 O).

Furthermore, to further validate the function of the selected amino acid metabolism-related signature, we chose GFPT2 as a hub gene and performed knockdown in the pancreatic cancer cell line BXPC3. Through transwell assays, we evaluated the impact of GFPT2 on invasion and migration in vitro. The results demonstrated that the knockdown of GFPT2 significantly decreased the migratory and invasive capabilities of pancreatic cancer cells, implying that GFPT2 is associated with malignant behavior and poor prognosis in pancreatic cancer (Fig. 11).

Discusssion

Pancreatic cancer represents a malignant neoplasm associated with a low survival rate, demonstrating rapidly increasing morbidity and mortality. Early metastasis and a lack of early diagnostic biomarkers often result in missed opportunities for curative resection at the time of diagnosis [21, 22]. Although significant advancements have been achieved in pancreatic cancer treatment, the therapeutic impact remains constrained, with mFOLFOX therapy and the combination of gemcitabine and paclitaxel being commonly utilized for patients with advanced pancreatic cancer and those necessitating postoperative adjuvant therapy [23, 24]. Furthermore, the efficacy of immunotherapy, which has proven effective in numerous gastrointestinal malignancies, is notably limited in pancreatic cancer [25]. The emergence of drug resistance to traditional adjuvant therapy, coupled with tumor microenvironment heterogeneity, can contribute to unfavorable therapeutic responses [5, 26, 27]. Thus, further exploration of the underlying mechanisms governing pancreatic cancer development and progression, along with an investigation into its relationship with alterations in the tumor immune microenvironment and drug resistance, is imperative. This endeavor will serve as a foundational basis for enhancing patient prognosis.

Intracellular metabolic reprogramming stands as a fundamental hallmark of cancer, enabling tumor cells to adapt to hypoxic environments and gain a survival advantage [28]. Emerging evidence underscores the vital role of amino acid metabolic reprogramming, in addition

(See figure on next page.)

Fig. 10 Validation of the expression levels in 4 prognostic genes. **A**, **B**, **C**, **D** Gene expression levels of GFPT2, P4HA1, HAAO and SAT2 in high- and low-risk groups in TCGA and ICGC cohorts. **E**, **F**, **G**, **H** Kaplan–Meier curves of GFPT2, P4HA1, HAAO and SAT2. **I**, **J**, **K**, **L** qPCR analysis identified the RNA expression levels of GFPT2, P4HA1, HAAO and SAT2 in HPNE, Panc1, T3M4 and CFPAC1 cell lines. **M**, **N** The protein expression levels of GFPT2,P4HA1,HAAO and SAT2 were detected by immunohistochemical staining in 40 pairs of PDAC tumor tissues and non-tumor tissues. **O** Correlation analysis of the 4 gene expression levels with prognosis in PDAC

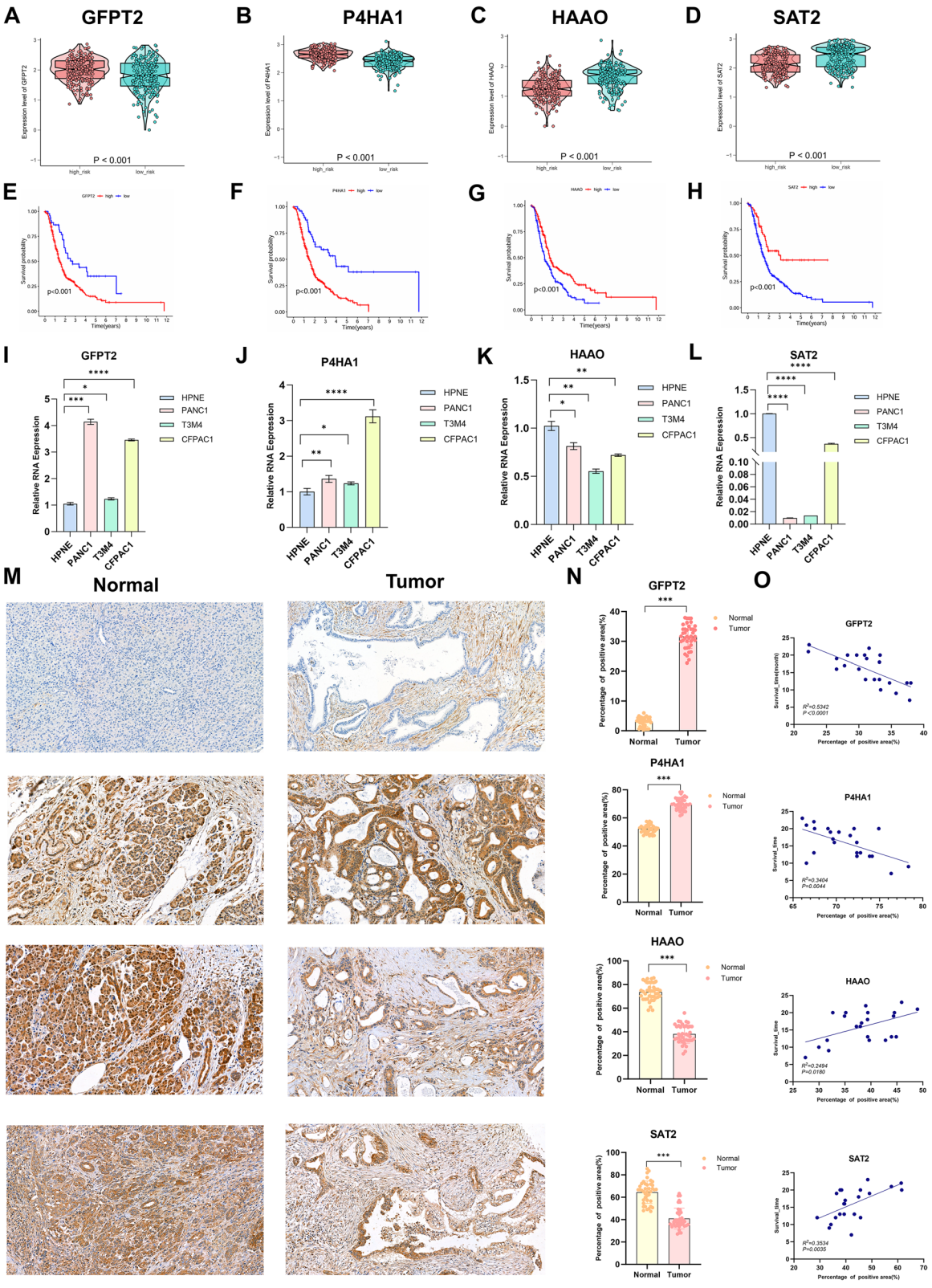

Fig. 10 (See legend on previous page.)

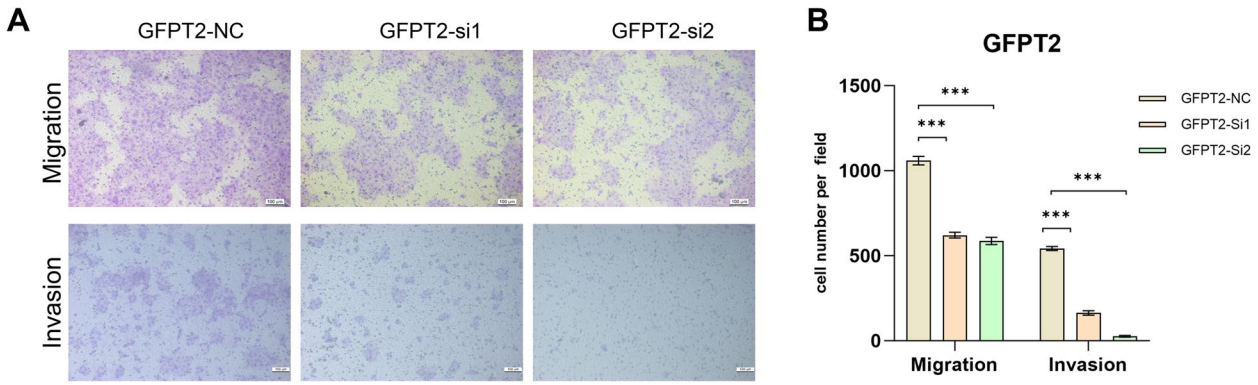

Fig. 11 Validation of the function of GFPT2. **A** Migration and invasion analyses of GFPT2 gene. **B** Knockdown of GFPT2 ameliorated the malignant ability of pancreatic cancer cells

to glucose and fatty acids, in tumorigenesis and cancer development [29–31]. An amino acid-enriched subtype has been identified by metabolomics analysis, highlighting the specific involvement of amino acids in PDAC [32]. Various amino acids and their metabolites exert diverse biological effects, such as supporting the tricarboxylic acid (TCA) cycle, maintaining cellular REDOX state, and participating in cell signal transduction [33–36]. Simultaneously, the interplay between amino acid metabolism, immune microenvironment, and epigenetic regulation also significantly influences tumor development [37, 38]. Leveraging amino acids as innovative focal points in the diagnosis, prognosis, and treatment of PDAC holds significant promise. The distinct amino acid metabolism of cancer cells, divergent from that of normal cells, renders it a viable therapeutic target.

In this study, we employed a combined approach of ssGSEA, WGCNA, and an integration of machine learning techniques to identify four amino acid metabolism related genes, including GFPT2, P4HA1, HAAO, and SAT2, demonstrating significant prognostic value in pancreatic cancer. Subsequently, we developed a prognostic model based on the 4 prognostic genes, and conducted validation in two internal and two external datasets to assess the model's ability of differentiation, calibration, and clinical benefit for the overall survival time of pancreatic cancer patients. Univariate and multivariate Cox regression analyses indicated that risk score of the novel model could serve as an independent prognosticator for pancreatic cancer (Supplementary Fig. 5). GFPT2 is a rate-limiting enzyme in the hexosamine biosynthesis pathway responsible for glycosylation, which enhances the glycosylation and nuclear translocation of p65, thereby activating the NF-κB pathway and promoting the progression and metastasis of CRC [39]. It has also been reported that the overexpression of GFPT2 and the heightened glutamine consumption in tumor cells can impede mitochondrial division in macrophages, thereby conferring resistance to macrophage phagocytosis [40]. GFPT2 knockdown can affect EMT, growth and invasion of breast cancer cells, and GFPT2 can also be used as a marker of oxidative stress and affect mitochondrial homeostasis [41]. Stomatin-like protein 2 induces pancreatic cancer metastasis by regulating GFPT2 expression under chemotherapy stress [42]. In addition, GFPT2 is also closely related to the tumorigenesis and progression of non-small cell lung cancer, gastric cancer, and ovarian cancer [43–45]. In conclusion, although the role of GFPT2 in pancreatic cancer was less reported, it has been correlated to tumor cell growth, invasion, and metastasis in a variety of tumors. We further demonstrated through in vitro functional assays that the knockdown of GFPT2 impedes the invasion and migration of pancreatic cancer cells, suggesting that GFPT2 expression is associated with the metastasis and malignant behavior of pancreatic cancer cells. P4HA1 is involved in proline metabolism and collagen synthesis, which has been implicated in the promotion of tumor progression across various cancers [46, 47]. Notably, P4HA1 exhibits high expression levels in PDAC and is correlated with a poor prognosis [48]. P4HA1 can also affect tumor cell glucose metabolism reprogramming and tumor dryness in pancreatic cancer [49]. Currently, limited research exists on the role of HAAO in cancer. HAAO pertains to the tryptophan metabolic pathway. However, available evidence suggests that the cancer-specific promoter hypermethylation of HAAO in prostate cancer is linked to risk stratification and it may serve as a prognostic marker [50]. SAT2 exhibits diamine N-acetyltransferase activity and participates in polyamine metabolism, primarily located in the extracellular exosomes. Currently, there is less research on SAT2 in tumors. Our study revealed a decreased expression of SAT2 in pancreatic cancer compared to normal tissues. Furthermore, high

SAT2 expression was related to improved overall survival and disease-free survival, suggesting its potential role as a protective factor in pancreatic cancer. To further elucidate the impact of amino acid metabolism-related genes on the prognosis of pancreatic cancer, it is necessary to conduct more comprehensive in vitro functional assays and in vivo studies. Given the complex biology of pancreatic ductal adenocarcinoma (PDAC), integrating these experimental models will aid in achieving a more nuanced understanding of how amino acid metabolism intersects with tumor growth, metastasis, and responses to therapy. These investigations will provide critical insights into the molecular mechanisms underlying the alterations in amino acid metabolism that may contribute to the aggressiveness of PDAC.

Additionally, we compared the amino acid-related signature with the C-index of six previously published prognostic models related to pancreatic cancer [51–56]. The results demonstrated that the amino acid-related signature exhibits superiority and robustness in predicting pancreatic cancer prognosis in both the training and validation cohorts (Supplementary Fig. 1). However, this study still has several limitations, including potential biases and the representativeness of the patient population. To enhance the appliance of our conclusions, we recommend that future research focuses on evaluating risk scores in a broader clinical cohort. Such studies should include prospective cohorts and multicenter collaborations to provide a comprehensive understanding of the reliability of risk scores for pancreatic cancer patients across different stages and treatment modalities. This approach will ultimately aid in refining personalized treatment strategies and improving prognostic accuracy in clinical practice.

We further delved into the correlation between the amino acid metabolism-related risk score and the immune microenvironment, immune evasion, and immunotherapy in pancreatic cancer. Our findings indicate that patients in the high-risk group are more prone to immune evasion and exhibit a poorer response to immunotherapy. Amino acid metabolism plays a crucial role in regulating immune cell function and shaping the immune microenvironment. The interplay between amino acid metabolism and immune responses has been widely documented. For instance, tryptophan metabolism through the kynurenine pathway can modulate T cell function and promote immune tolerance [57]. Additionally, arginine metabolism influences T cell proliferation and phenotype, impacting the anti-tumor immune response [58]. Furthermore, the availability of amino acids like glutamine impacts the activation and function of immune cells [59, 60]. Amino acid metabolism also contributes to the regulation of immune checkpoints and the tumor immune microenvironment [61]. Dysregulated amino acid metabolism in cancer can lead to the accumulation of immunosuppressive metabolites, creating an immune-suppressive microenvironment. Understanding the intricate relationship between amino acid metabolism and immune function is essential for developing novel immunotherapeutic strategies. Targeting amino acid metabolic pathways may offer promising avenues for modulating the immune response and improving the efficacy of immunotherapies in cancer. Further investigation in this area is warranted to elucidate the specific mechanisms and identify potential therapeutic targets.

In the analysis of immune checkpoints, it is essential not only to emphasize the importance of longitudinal studies and functional assays but also to explore how these expression changes specifically impact treatment outcomes over time. By incorporating longitudinal analyses, we can identify not only the immediate effects of therapeutic interventions but also the long-term consequences of gene expression alterations on tumor behavior and the activation of immune responses. Future investigations may include the utilization of patient-derived xenograft models to observe the dynamic interactions between tumor cells and the immune microenvironment in real-time. Furthermore, advanced methodologies such as single-cell RNA sequencing can provide deeper insights into the heterogeneity of immune cell populations and their functional states both before and after treatment. Integrating transcriptomic data with proteomic and metabolomic analyses will enable a more comprehensive understanding of how the tumor microenvironment influences treatment efficacy. By employing these multifaceted approaches, we aim to identify reliable biomarkers that accurately predict patient responses, thereby facilitating the customization of therapeutic strategies in a more personalized manner.

Moreover, translating the amino acid metabolism relate signature into clinical practice constitutes a significant and intricate undertaking. First and foremost, it is imperative to validate this signature across multiple independent patient cohorts to ascertain its generalizability and consistency. Following this validation, prospective studies should be undertaken to assess the practical utility of the gene signature in prognostic evaluations. Additionally, it is vital to explore how this signature may impact clinical decision-making, particularly regarding the selection of treatment protocols. By adopting this comprehensive strategy, we aspire to effectively translate our research findings into a clinically relevant prognostic tool, ultimately facilitating more personalized treatment options for patients.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12885-024-13374-4.

Supplementary Material 1.

Supplementary Material 2.

Acknowledgements
We acknowledge the funding supported by the National Natural Science Foundation of China (No. 81972321 and 82273455 to L. Y., 82303504 to B.R.), Non-proft Central Research Institute Fund of Chinese Academy of Medical Sciences (No. 2018PT32014), National Multidisciplinary Cooperative Diagnosis and Treatment Capacity Building Project for Major Diseases (No. Z190022), the CAMS Innovation Fund for Medical Sciences (CIFMS) (No. 2021-I2M-1-002 to Y. Z.), China Postdoctoral Science Foundation(No. 2024T170071), and National High Level Hospital Clinical Research Funding (No. 2022-PUMCH-D-001).

Authors' contributions
B.R., L.Y. and Y. Z. designed and revised the manuscript. X. L., X. W. and J.R. conducted the bioinformatic analyses and wrote the manuscript. M. G., D. J., Y. T. and J. B. did statistical analysis and elaborated the figures. Y. F., F.Z., R. X., and X. L. polished the manuscript and gave useful suggestions. All authors reviewed the manuscript.

Data availability
The datasets utilized in the present study are accessible in public repositories including TCGA, GEO, and ICGC. The source code is available upon reasonable request from the corresponding author.

Declarations

Ethics approval and consent to participate
The investigation involving human participants was reviewed and approved by the Medical Ethics Committee of Peking Union Medical College Hospital. Written informed consent was obtained from all patients before their participation in the study.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Department of General Surgery, Peking Union Medical College Hospital, Peking Union Medical College, Chinese Academy of Medical Sciences, Peking, Beijing 100023, People's Republic of China. [2]Key Laboratory of Research in Pancreatic Tumor, Chinese Academy of Medical Sciences, Beijing 100023, People's Republic of China. [3]National Science and Technology Key Infrastructure On Translational Medicine in Peking Union Medical College Hospital, Beijing 100023, People's Republic of China.

Received: 10 June 2024 Accepted: 19 December 2024
Published online: 03 January 2025

References

1. Klein AP. Pancreatic cancer epidemiology: understanding the role of lifestyle and inherited risk factors. Nat Rev Gastroenterol Hepatol. 2021;18(7):493–502.
2. Halbrook CJ, Lyssiotis CA. Pasca di Magliano M, Maitra A. Pancreatic cancer: Advances and challenges Cell. 2023;186(8):1729–54.
3. Kleeff J, Korc M, Apte M, et al. Pancreatic cancer Nature reviews Disease primers. 2016;2:16022.
4. Dbouk M, Katona BW, Brand RE, et al. The Multicenter Cancer of Pancreas Screening Study: Impact on Stage and Survival. Journal of clinical oncology : official journal of the American Society of Clinical Oncology. 2022;40(28):3257–66.
5. Sherman MH, Beatty GL. Tumor Microenvironment in Pancreatic Cancer Pathogenesis and Therapeutic Resistance. Annu Rev Pathol. 2023;18:123–48.
6. Zhang S, Fang W, Zhou S, et al. Single cell transcriptomic analyses implicate an immunosuppressive tumor microenvironment in pancreatic cancer liver metastasis. Nat Commun. 2023;14(1):5123.
7. Vettore L, Westbrook RL, Tennant DA. New aspects of amino acid metabolism in cancer. Br J Cancer. 2020;122(2):150–6.
8. Liu N, Shi F, Yang L, Liao W, Cao Y. Oncogenic viral infection and amino acid metabolism in cancer progression: Molecular insights and clinical implications. Biochim Biophys Acta. 2022;1877(3): 188724.
9. Sivanand S, Vander Heiden MG. Emerging Roles for Branched-Chain Amino Acid Metabolism in Cancer. Cancer Cell. 2020;37(2):147–56.
10. Yang L, Chu Z, Liu M, et al. Amino acid metabolism in immune cells: essential regulators of the effector functions, and promising opportunities to enhance cancer immunotherapy. J Hematol Oncol. 2023;16(1):59.
11. Yoo HC, Han JM. Amino acid metabolism in cancer drug resistance. Cells. 2022;11(1):140.
12. Wise DR, Thompson CB. Glutamine addiction: a new therapeutic target in cancer. Trends Biochem Sci. 2010;35(8):427–33.
13. Lind DS. Arginine and cancer. J Nutr. 2004;134(10 Suppl):2837S-41S discussion 53S.
14. Yang M, Vousden KH. Serine and one-carbon metabolism in cancer. Nat Rev Cancer. 2016;16(10):650–62.
15. Peng H, Wang Y, Luo W. Multifaceted role of branched-chain amino acid metabolism in cancer. Oncogene. 2020;39(44):6747–56.
16. Issa NT, Stathias V, Schürer S, Dakshanamurthy S. Machine and deep learning approaches for cancer drug repurposing. Semin Cancer Biol. 2021;68:132–42.
17. Tran KA, Kondrashova O, Bradley A, Williams ED, Pearson JV, Waddell N. Deep learning in cancer diagnosis, prognosis and treatment selection. Genome medicine. 2021;13(1):152.
18. Morabito S, Reese F, Rahimzadeh N, Miyoshi E, Swarup V. hdWGCNA identifies co-expression networks in high-dimensional transcriptomics data. Cell reports methods. 2023;3(6): 100498.
19. Liu Z, Liu L, Weng S, et al. Machine learning-based integration develops an immune-derived lncRNA signature for improving outcomes in colorectal cancer. Nat Commun. 2022;13(1):816.
20. Liu X, Ren B, Fang Y, et al. Comprehensive analysis of bulk and single-cell transcriptomic data reveals a novel signature associated with endoplasmic reticulum stress, lipid metabolism, and liver metastasis in pancreatic cancer. J Transl Med. 2024;22(1):393.
21. Jiang S, Fagman JB, Ma Y, et al. A comprehensive review of pancreatic cancer and its therapeutic challenges. Aging. 2022;14(18):7635–49.
22. Greenhalf W, Thomas A. Combination therapy for the treatment of pancreatic cancer. Anticancer Agents Med Chem. 2011;11(5):418–26.
23. Conroy T, Desseigne F, Ychou M, et al. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. N Engl J Med. 2011;364(19):1817–25.
24. Von Hoff DD, Ervin T, Arena FP, et al. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. N Engl J Med. 2013;369(18):1691–703.
25. Mukherji R, Debnath D, Hartley ML, Noel MS. The Role of Immunotherapy in Pancreatic Cancer. Current oncology (Toronto, Ont). 2022;29(10):6864–92.
26. Ren B, Cui M, Yang G, et al. Tumor microenvironment participates in metastasis of pancreatic cancer. Mol Cancer. 2018;17(1):108.
27. Connor AA, Gallinger S. Pancreatic cancer evolution and heterogeneity: integrating omics and clinical data. Nat Rev Cancer. 2022;22(3):131–42.
28. Hanahan D. Hallmarks of Cancer: New Dimensions. Cancer Discov. 2022;12(1):31–46.
29. Locasale JW. Serine, glycine and one-carbon units: cancer metabolism in full circle. Nat Rev Cancer. 2013;13(8):572–83.
30. Lei MZ, Li XX, Zhang Y, et al. Acetylation promotes BCAT2 degradation to suppress BCAA catabolism and pancreatic cancer growth. Signal Transduct Target Ther. 2020;5(1):70.
31. Wang Z, Yip LY, Lee JHJ, et al. Methionine is a metabolic dependency of tumor-initiating cells. Nat Med. 2019;25(5):825–37.

32. Guo P, Teng T, Liu W, et al. Metabolomic analyses redefine the biological classification of pancreatic cancer and correlate with clinical outcomes. Int J Cancer. 2022;151(10):1835–46.
33. Jeong SM, Hwang S, Park K, Yang S, Seong RH. Enhanced mitochondrial glutamine anaplerosis suppresses pancreatic cancer growth through autophagy inhibition. Sci Rep. 2016;6:30767.
34. Araújo WL, Martins AO, Fernie AR, Tohge T. 2-Oxoglutarate: linking TCA cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. Front Plant Sci. 2014;5:552.
35. Koppula P, Zhang Y, Zhuang L, Gan B. Amino acid transporter SLC7A11/xCT at the crossroads of regulating redox homeostasis and nutrient dependency of cancer. Cancer communications (London, England). 2018;38(1):12.
36. Meijer AJ, Lorin S, Blommaart EF, Codogno P. Regulation of autophagy by amino acids and MTOR-dependent signal transduction. Amino Acids. 2015;47(10):2037–63.
37. Sullivan MR, Danai LV, Lewis CA, et al. Quantification of microenvironmental metabolites in murine cancers reveals determinants of tumor nutrient availability. eLife. 2019;8:e44235.
38. Qian L, Li N, Lu XC, et al. Enhanced BCAT1 activity and BCAA metabolism promotes RhoC activity in cancer progression. Nat Metab. 2023;5(7):1159–73.
39. Liu L, Pan Y, Ren X, et al. GFPT2 promotes metastasis and forms a positive feedback loop with p65 in colorectal cancer. Am J Cancer Res. 2020;10(8):2510–22.
40. Li J, Ye Y, Liu Z, et al. Macrophage mitochondrial fission improves cancer cell phagocytosis induced by therapeutic antibodies and is impaired by glutamine competition. Nature cancer. 2022;3(4):453–70.
41. Wang Q, Karvelsson ST, Kotronoulas A, Gudjonsson T, Halldorsson S, Rolfsson O. Glutamine-Fructose-6-Phosphate Transaminase 2 (GFPT2) Is Upregulated in Breast Epithelial-Mesenchymal Transition and Responds to Oxidative Stress. Molecular & cellular proteomics : MCP. 2022;21(2):100185.
42. Chao D, Ariake K, Sato S, et al. Stomatin-like protein 2 induces metastasis by regulating the expression of a rate-limiting enzyme of the hexosamine biosynthetic pathway in pancreatic cancer. Oncol Rep. 2021;45(6):90.
43. Szymura SJ, Zaemes JP, Allison DF, et al. NF-κB upregulates glutamine-fructose-6-phosphate transaminase 2 to promote migration in non-small cell lung cancer. Cell Commun Signal. 2019;17(1):24.
44. Yang S, Li G, Yin X, et al. Cancer-associated fibroblast expression of glutamine fructose-6-phosphate aminotransferase 2 (GFPT2) is a prognostic marker in gastric cancer. The journal of pathology Clinical research. 2023;9(5):391–408.
45. Zhou L, Luo M, Cheng LJ, Li RN, Liu B, Linghu H. Glutamine-fructose-6-phosphate transaminase 2 (GFPT2) promotes the EMT of serous ovarian cancer by activating the hexosamine biosynthetic pathway to increase the nuclear location of β-catenin. Pathol Res Pract. 2019;215(12): 152681.
46. Kappler M, Kotrba J, Kaune T, et al. P4HA1: A single-gene surrogate of hypoxia signatures in oral squamous cell carcinoma patients. Clinical and translational radiation oncology. 2017;5:6–11.
47. Atkinson A, Renziehausen A, Wang H, et al. Collagen Prolyl Hydroxylases Are Bifunctional Growth Regulators in Melanoma. J Invest Dermatol. 2019;139(5):1118–26.
48. Cao XP, Cao Y, Li WJ, Zhang HH, Zhu ZM. P4HA1/HIF1α feedback loop drives the glycolytic and malignant phenotypes of pancreatic cancer. Biochem Biophys Res Commun. 2019;516(3):606–12.
49. Cao X, Cao Y, Zhao H, Wang P, Zhu Z. Prolyl 4-hydroxylase P4HA1 Mediates the Interplay Between Glucose Metabolism and Stemness in Pancreatic Cancer Cells. Curr Stem Cell Res Ther. 2023;18(5):712–9.
50. Li Y, Meng L, Shi T, Ren J, Deng Q. Diagnosis and prognosis potential of four gene promoter hypermethylation in prostate cancer. Cell Biol Int. 2021;45(1):117–26.
51. Huo J, Wu L, Zang Y. Development and Validation of a Novel Metabolic-Related Signature Predicting Overall Survival for Pancreatic Cancer. Front Genet. 2021;12: 561254.
52. Tian X, Zheng J, Mou W, et al. Development and validation of a hypoxia-stemness-based prognostic signature in pancreatic adenocarcinoma. Front Pharmacol. 2022;13: 939542.
53. Mao M, Ling H, Lin Y, Chen Y, Xu B, Zheng R. Construction and Validation of an Immune-Based Prognostic Model for Pancreatic Adenocarcinoma Based on Public Databases. Front Genet. 2021;12: 702102.
54. Xu Y, Li H, Lan A, et al. Cuprotosis-Related Genes: Predicting Prognosis and Immunotherapy Sensitivity in Pancreatic Cancer Patients. Journal of oncology. 2022;2022:2363043.
55. Xu F, Zhang Z, Zhao Y, Zhou Y, Pei H, Bai L. Bioinformatic mining and validation of the effects of ferroptosis regulators on the prognosis and progression of pancreatic adenocarcinoma. Gene. 2021;795: 145804.
56. Xu D, Wang Y, Liu X, et al. Development and clinical validation of a novel 9-gene prognostic model based on multi-omics in pancreatic adenocarcinoma. Pharmacol Res. 2021;164: 105370.
57. Stone TW, Williams RO. Modulation of T cells by tryptophan metabolites in the kynurenine pathway. Trends Pharmacol Sci. 2023;44(7):442–56.
58. Geiger R, Rieckmann JC, Wolf T, et al. L-Arginine Modulates T Cell Metabolism and Enhances Survival and Anti-tumor Activity. Cell. 2016;167(3):829-42.e13.
59. Yu Q, Tu H, Yin X, et al. Targeting Glutamine Metabolism Ameliorates Autoimmune Hepatitis via Inhibiting T Cell Activation and Differentiation. Front Immunol. 2022;13: 880262.
60. Ma EH, Bantug G, Griss T, et al. Serine Is an Essential Metabolite for Effector T Cell Expansion. Cell Metab. 2017;25(2):345–57.
61. Wang W, Zou W. Amino Acids and Their Transporters in T Cell Immunity and Cancer Therapy. Mol Cell. 2020;80(3):384–95.

Publisher's Note