An Accurate Prognostic Model and Treatment for Pancreatic Carcinoma Based on the Tumor Immune Microenvironment

Shujie Wang

Southern Medical University

zhenchong li

South China University of Technology School of Medicine

chunsheng liu

Shantou University Medical College

qi zhou

Sun Yat-sen University First Affiliated Hospital

zuyi ma

Shantou University Medical College

zixuan zhou

South China University of Technology

Hongkai Zhuang

Shantou University Medical College

Shanzhou Huang

Southern Medical University

Chanzhao Zhang

Southern Medical University

Baohua Hou (✉ hbh1000@126.com)

Southern Medical University <https://orcid.org/0000-0002-4239-1367>

Research Article

Keywords: pancreatic adenocarcinoma, immune microenvironment, immune classification, immune subtypes, prognosis

Posted Date: December 6th, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-1121256/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background

Pancreatic adenocarcinoma (PAAD) is a highly malignant cancer with a poor prognosis. The tumor microenvironment (TME) is closely related to tumorigenesis, progression, and treatment. However, the relationship between TME immune cell genes and prognosis in PAAD is currently unclear.

Methods

In this study, we identified three prognostic subtypes based on the TME by using data from The Cancer Genome Atlas (TCGA) database, The International Cancer Genome Consortium (ICGC) database and University of California Santa Cruz (UCSC) database. The Silhouette plot analysis was used to evaluate 758 immune genes expression in PAAD from each database, then to divide all samples into three subtypes (Clusters A, B, C) by Lasso's binomial logistic regression. We analyzed the relationship between subtypes and prognosis by the survival R package. CIBERSORT was used for evaluating the expression changes of immune cells. We detect the copy number variation areas between two groups through GISTIC 2.0 algorithm. The TIDE network tool was used to predict the response of immune therapy.

Results

We defined three clusters (Clusters A, B, and C) based on the analysis of immune gene expression. Cluster B got a worse prognosis than the other two clusters. The Cluster B group had the highest level of Macrophages M0 and Macrophage M2. NK cell resting was much higher in Cluster B than other groups in TME. Gene KRAS was mutated in 77% of all samples. Cluster C had a better immune therapy effect than others.

Conclusions

We found a new model to predict patients' prognosis who with pancreatic adenocarcinoma. Cluster B had the significant worse prognosis than other groups. Patients in Cluster C could get better treatment effect by using immunotherapy.

Introduction

Pancreatic adenocarcinoma (PAAD) is one of the most fatal malignancies in the world, representing 2.5% of all cancers, as well as one of the most difficult to diagnose and cure[1]. Almost 85% of PAADs are pancreatic ductal adenocarcinomas (PDAC), which are originating from the glandular epithelium[2]. PAAD has a 5-year survival rate of 9%, making it one of malignant tumors with poorest prognosis[3]. The early diagnosis rate of pancreatic cancer is very low, for 80-90% of patients have unresectable tumors at the moment of diagnosis[3]. Even though the surgical rate is increasing, the cure rate is low and the recurrent rate remains very high. The median overall survival varies between 24–30 months[4]. The clinical

manifestations of PAAD include the cancer site, the course of the disease, the presence of metastasis, and the involvement of adjacent organs[5, 6].

The major challenges in the treatment of pancreatic cancer are as follows: the metastasis and immunosuppressive tumor microenvironment and the difficulty in treating PAAD may be largely due to the protective stromal barriers that effectively limit drug delivery, and the distribution of immune cell populations[7-10]. There is no doubt that stroma is an important feature of this disease, regulating central processes such as tumor growth, angiogenesis, drug response, and metastasis [8]. As such, the tumor microenvironment influences cancer initiation and progression, making it a target for drug development efforts[11, 12]. Therefore, it is imperative to better understand the pathogenesis of the tumor and explore more treatment strategies to improve the prognosis of PAAD patients. Malignant solid tumor tissues are composed of tumor cells, tumor-associated normal epithelium cells, vascular cells, immune cells, and stromal cells[13]. Tumor-associated fibroblasts, adipocytes, pericytes, mesenchymal stem cells, endothelial cells, lymphocytes, and extracellular matrix are components of the tumor microenvironment[14]. A noticeable feature of pancreatic cancer is an extensive desmoplastic stroma comprised of fibroblasts, extracellular matrix (ECM), and immune cells[15-17]. There were four subclasses in the stroma functionally: (i) structurally vascularized, (ii) activated, (iii) inflammatory, and (iv) immune [18]. According to the relevant research, the immune microenvironment surrounding cancer cells can recognize and restrain tumor growth or promote progression [19, 20].

However, the infiltration of immune cells in the tumor environment of PAAD is poorly understood, where researchers assessed each cell type individually but did not consider the immune microenvironment as a whole[21-23]. Thus we need more efforts to specify the role and clinical relevance of the immune contexture in PAAD. To identifying the quality and quantity of tumor-site immune responses is critical because it may help determine which patients can benefit from immunotherapy and improve our understanding of tumor host biology. In this study, we intended to build a prediction model of poor prognosis for PAAD and find the genes or immune cells related to the immune microenvironment leading to inferior prognosis. We found clinically relevant immune clusters with progressive immune infiltration. By characterizing the immune components of the cluster, we identified a group of tumorigenic immune infiltrates associated with poor prognosis. Further phenotypic analysis revealed two mutually exclusive aggressive tumor phenotypes in PAAD. Then we tried to find the possibility of insight therapeutic targets for PAAD.

Methods

Data collection and processing

Expression data and prognostic clinical data of PAAD were obtained from University of California Santa Cruz (UCSC) database. Mutation data were obtained from The Cancer Genome Atlas (TCGA) database. Copy number variation (CNV) data were obtained from Broad Institute Genome Data Analysis Center (GDAC) Firehose database. Expression profile data of pancreatic cancer (PACA) with survival information

were obtained from The International Cancer Genome Consortium (ICGC) database. All collected samples from each database have been shown in Table 1.

Unsupervised clustering to obtain immune clusters

The expression profiles of tumor samples from TCGA-PAAD and PACA-CA datasets were extracted for subsequent analysis. The list of immune genes was downloaded from nCounter® PanCancer Immune Profiling Panel (Human), which were, respectively, intersected with the expression profiles of TCGA-PAAD and PACA-CA tumor samples to obtain the expression profiles of the training expression profile and the verification expression profile of 758 immune genes. According to the immune gene expression profile obtained in the previous step, we evaluated the optimal class number K in combination with the Silhouette algorithm, and then conducted hierarchical clustering for all tumor samples.

CIBERSORT analysis

CIBERSORT (<http://cibersort.stanford.edu/>) is a deconvolution algorithm based on gene expression that can evaluate the expression changes of a gene set relative to all other genes in the sample [24]. Thus, the concentration of tumor-infiltrating immune cells (TIICs) can be accurately estimated by this process. The continued performance of CIBERSORT has led to increased interest in cell heterogeneity studies [22, 23]. Our current analysis assessed the immune response of 22 TIICs in PAAD using CIBERSORT to assess their correlation with survival [24]. In short, a dataset of gene expression is created using a standard annotation file and uploaded to the CIBERSORT website portal, where the algorithm uses its default signature matrix for a thousand permutations. CIBERSORT estimates the P value of deconvolution by Monte Carlo sampling and establishes the confidence of the result. We divided PACA-CA database and TCGA-PAAD database, including 139 samples, into two clusters and then used the data to make plots.

Comparison of Survival Prognosis

Among the immune subtypes with survival data, the relationship between the immune subtypes and the overall survival of patients was analyzed using the survival R packet. The survival difference was tested by log-rank, and $P < 0.05$ was considered statistically significant. The Kaplan-Meier curve shows the survival differences between different immune subtypes. An R-Package "Estimate" was then used to evaluate stromal score, immune score, and tumor purity based on immune gene expression profiles.

Binomial logistic regression

We designed a classification model to predict which sample belonged to the poor prognosis group by using lasso's binomial logistic regression. At the meantime, a key gene was obtained, and the Risk Score of each sample was calculated according to the weight of these genes, that is, the expression value of each key gene multiplied by its weight and summed up.

Gene set enrichment analysis (GSEA) analysis

Gene set collections for canonical pathways (curated gene sets, 1329 gene sets) were downloaded from the Molecular Signatures Database version 6.2 (www.broadinstitute.org/gsea/msigdb/collections.jsp). Gene set enrichment scores were calculated with GSEA package version 1.32.0 with RNA-seq parameters [28]. Differential gene collection enrichment was determined by LIMMA packets. A statistically significant threshold was recorded in the results. The DESeq2 packet of R was used to analyze the differential expression between Cluster B and Cluster A & C groups, according to the threshold $FDR < 0.05$ and $|\log_2FC| > 1$. Next, GSEA function in R's cluster Profiler package was used to analyze the differently expressed genes in the curated gene sets of GSEA.

Copy Number Variation (CNV)

The CNV data of Sequence Tagged S (STS) samples were downloaded from FireHose website (<http://gdac.broadinstitute.org/>), which stored integrated data sourced from TCGA database. The corresponding GISTIC_2.0 module in Gene Pattern was used to detect the copy number variation areas in tumor samples from different groups using its GISTIC algorithm.

Tumor Immune Dysfunction and Exclusion (TIDE)

TIDE (<http://tide.dfci.harvard.edu>) was used to identify factors that underlie tumor immune escape. TIDE combines and models data from 189 human cancer studies, including a total of 33,197 samples. We hypothesize and verify that an accurate genetic marker that mimics tumor immune escape can serve as a reliable alternative biomarker for predicting immune checkpoint blockade (ICB) responses. The TIDE network tool was used to predict the response of different tumor samples to immunotherapy based on the gene expression data of tumor samples.

Statistical analysis

All statistical analyses were performed using IBM SPSS Statistics for Windows, Version 22 (IBM Corp. Armonk, NY, USA) and GraphPad Prism 7.0 (GraphPad Software, La Jolla, CA, USA). The chi-square test was used to analyze the relationship between PDAC and clinicopathological characteristics. Student's-t test was used to analyze the differences among the three groups. Kaplan-Meier method was used to establish survival curves and log-rank test was used to compare survival differences. All statistical results with $P < 0.05$ were considered statistically significant.

Results

Cluster of immune genes in PDAC from TCGA-PAAD. The expression data of 178 samples from TCGA-PAAD were measured by the nCounter® PanCancer Immune Profiling array, an immune gene array. Silhouette plot analysis from 3 to 10 clusters indicated that 3 clusters demonstrated the best segmentation of the nCounter (**Fig 1A**), showing it as a better pick than more numerous clusters. Then all samples were clustered by the hierarchical clustering method (**Fig 1B**). Finally, we got three clusters:

Cluster A, Cluster B and Cluster C (**Fig 1C, Supplementary Table 1**). We concluded that Clusters A–C reflect gradual immune infiltration and were therefore called immune clusters.

Characterization and prognosis of Clusters

We evaluated the stemness features of the 3 clusters based on the methods reported by previous study [29] and compared them by the Wilcoxon Signed Rank Test. We found fundamental differences in stemness scores among the 3 clusters (**Fig 2A, Supplementary Table 2**). Cluster C had the highest stemness feature score. The scatter graph with bars showed that Cluster A had the highest stromal scores, immune scores, and tumor purity(**Supplementary Table 3**), while Cluster C got the lowest score (**Fig 2B-D**). There were significant differences in stromal score, immune score and tumor purity among Cluster A vs Cluster B, Cluster A vs Cluster C and Cluster B vs Cluster C.

We designed a classification model to predict which sample belonged to the poor prognosis group by using lasso's binomial logistic regression. At meantime, we got 10 genes, which include CCL4, CD8A, EOMES, FLT3, ITGB4, MUC1, PECAM1, STAT5B, TLR7 AND TNFRSF8(**Supplementary Table 4**), from the model and used it to calculate the tumor scores(**Supplementary Table 5**). The patient had higher scores which proved the worse prognosis. The clinical survival information revealed that the prognosis of the three subtypes had a significant differences (**Fig 2E**). After combing Cluster A and Cluster C to group Cluster A & C, we found Cluster B (with intermediate levels of immune infiltration) associated with a worse survival prognosis than Cluster A & C (**Fig 2F**), indicating Cluster B as a group predicting inferior survival outcomes. Moreover, Cluster B got higher risk scores compared to Cluster A & C group based on the key genes related to inferior survival (**Fig 2G**). Then we divided the samples into groups according to some clinical features, such as smoking status and pancreatic tumor's site. It was found that the risk scores of Cluster B and Cluster A & C groups were significantly higher in smoking patients (**Fig 2H**). And the tumor in Cluster B which had higher cancer scores compared to Cluster A & C in various pancreatic tumor's sites (**Fig 2I**).

Verification of the model via ICGC data

We downloaded the PACA-CA data from ICGC database, then divided all samples into three clusters using the Silhouette Coefficient. Then we used the ICGC database to certify the model's effect. In ICGC database, Cluster B group got higher risk scores compared to the other groups in all cancer samples (**Fig 3A**). Then divided the samples into various groups according to some clinical features. We could find that the Cluster B got higher scores no matter whether the patient smoked or not, nor the patient were diagnosed with diabetes or not. (**Fig 3B and G**). The location of the tumor in the pancreas does not influence the worse prognosis in Cluster B (**Fig 3C**). In tumor characters, such as tumor with or without metastatic tumor stage and tumor's dimension, Cluster B got a higher score than other clusters (**Fig 3D-F**). The survival curve showed a significant differences in prognosis among Cluster A, Cluster B, and Cluster C (**Fig 4A-C**). Cluster B showed a worse survival outcome when compared to the combined group Cluster A & C (**P<0.033, Fig 4D**). The model showed a strong prediction performance demonstrated by the

ROC curve in 10-year follow-up with an AUC=0.645. The results proved that the model has outstanding prediction power in PAAD.

The high expression of NK cells resting in TME predicted worse prognosis in Cluster B

We used the CIBERSORT algorithm to evaluate the immune microenvironment composition of PAAD samples. For each group, we calculated the median of CIBERSORT absolute scores for each of the 22 cell types in each cohort (**Supplementary Table 6**). The level of macrophages M0 and other immune cells had significant differences between Cluster B and Cluster A & C (**Figs 5A and B**). Macrophage M0 (**P<0.001**) showed an apparent upward trend in the infiltration of immune cells in Cluster B. We found that the Cluster B group had the highest level of Macrophages M0 and Macrophage M2, whereas Cluster A & C group had lower levels of these cell types. The immune cells in the tumor samples were significantly correlated with the prognosis, 55 differentially expressed genes were screened (**Fig 5C**). The different NK cells resting expression in TME showed significant difference in prognosis (**P=0.024**), while Cluster B had higher NK cell resting expression (**P<0.05, Fig 5**). Therefore we supposed that NK cell resting may be the key point of affecting the patient's prognosis in PAAD.

Phenotypic analysis of the immune clusters

To further characterize the phenotype associated with the poor prognosis in Cluster B, we identified the genes significantly overexpressed in Cluster B. We found 55 genes upregulated in Cluster B when compared to Cluster A & C (**Table 2, Supplementary Table 7**). These genes are involved in stem cell biology and EMT. To further characterize the relationship between immune clusters and cancer cell phenotypes, we used gene sets associated with EMT, stem cells, hypoxia, and proliferation. There were a total of seven gene sets from msigDB and one EMT-related signature from Tan et al.[30] were selected. We used the GSVA method to calculate the average enrichment fraction of gene sets for each cluster and cohort. This score reflects the activity of each pathway/gene set in the immune cluster [30]. These genes are mainly related to "BREAST_CANCER_NORMAL_LIKE_UP, REACTOME_CANCER_LUMINAL_B_DN, SANSOM_APC_TARGETS_DN, POOLA_INVASIVE_BREAST_CANCER_UP, WALLACE_PROSTATE_CANCER_RACE_UP, REACTOME_IMMUNE_SYSTEM" (**Fig 6A, Supplementary Table 8**). To formally identify which gene set scores explained Cluster B, we tested how each gene set contributed to Cluster B vs Clusters A&C using generalized linear models (**Fig 6B, Supplementary Table 9**). To further characterize the relationship between gene signatures and survival, we performed multivariable analysis, which indicated that "cell motility, mitotic cell cycle, hypoxia, mammary stem cell" are related to poor prognosis (**Fig 6C**). These results suggested there was a certain association between immune clusters with gene signatures.

Gene mutations and immunotherapy response prediction in different subtypes

R's Maftools package was used to analyze the highly mutated genes in tumor samples in different groups (**Fig 7A**). KRAS was mutated in 77% of the samples, followed by TP53 and SMAD4, which were mutated in 64% and 24% of the samples, respectively. We found that Cluster B and Cluster A & C had

frequent amplification on chromosomes 8q and 18q, and frequent deletion on chromosome 9q via GISTIC_2.0 of R's GenePattern package (Figs 7B and C). The TIDE network tool was used to predict the response of different tumor samples to immunotherapy based on the gene expression data of tumor samples. Among them (Supplementary Table 10), a high TIDE prediction score was associated with poor immune checkpoint inhibition and a poor prognosis. There was no significant difference in TIDE scores between the Cluster B and Cluster A&C groups (Fig 8A). However, Cluster C had a better immune therapy effect than the other two clusters (Fig 8B).

Discussion

Tumor microenvironment plays an important role in the pathogenesis of PAAD. In this study, we present a novel immune-related subtype that is associated with PAAD prognosis. The immune cluster described here depends on the abundance and composition of the immune cell infiltrates. We used an established computational resource (CIBERSORT) to explore the gene expression profiles of downloaded samples to infer the density of 22 types of immune cells. The CIBERSORT algorithm applied initially to the 22 immune cell subtypes helped assess their differing concentrations in Cluster B and Cluster A & C groups for prognosis.

After the precise grouping, we got three clusters in TCGA-PAAD, Cluster A, Cluster B and Cluster C. Then we found the patients in Cluster C got better prognosis than the other two clusters, so we predicted those patients with the higher stem-like feature and lower immune scores which had prolonged survival. Cluster B had the significant worse prognosis than other groups. Because the purpose of this study was to predict the prognosis of patients through the tumor immune microenvironment, the reanalysis of the prognosis showed that the prognosis of Group B still had the worst prognosis after combining the Cluster A and Cluster C.

The investigation of tumor-infiltrating immune cells (TIICs) has greatly improved our knowledge of the nature of tumor-immune interactions. The presence of TIICs has been associated with a favorable prognosis across multiple cancer type [32-34]. We found that NK cells resting was higher in Cluster B than in Cluster A & C group. And the NK cells resting had significantly affected the prognosis of patients who diagnosed as pancreatic adenocarcinoma among all immune cells. Recent evidences suggest that NK cells, besides their direct cytolytic effect against tumor cells, may also shape the adaptive immune response toward a T helper 1 (Th1) profile, which is thought to favor antitumor responses [35, 36]. Despite the complexity of NK cell receptors, NK cells appear to be a valuable and potentially more versatile alternative T cell in antitumor immunotherapy. The NK cells do not need to select or amplify specific clones to achieve antitumor activity. More importantly, they can be safely used for adoptive immunotherapy in allogeneic Settings [37, 38]. According to this study, we found that high resting NK cells infiltrating in the tumor microenvironment had worse prognosis in Cluster B.

In the present study, we found the key genes related to pancreatic cancer immunity. The score of Cluster B was higher than that of the other groups, indicating that the group with higher score had a poorer

prognosis (Fig 2C). Correlating with any influence factors, the score of group B was higher and the prognosis was worse (Fig 2D-H). We verified this classification method through PACA-CA data, which also testifying the prognosis of Cluster B was worse. The ROC curve was satisfying, which proved that this classification method was reliable and feasible. However, there was no significant difference in some influencing factors among the poor prognosis group of PACA-CA data species. It may suggest that the immune microenvironment may have little correlation with these influencing factors. It indicating that the gene signatures such as EMT, stem cell, and others, which may have correlation with PAAD. The Forest plot shows the risk degree of each pathway in Figure 7C, which indicates that the pathways like STEM_CELL_PATHWAY are risk factors for pancreatic cancer survival, and it is speculated that it may be related to the classification of Cluster B with poor prognosis.

According to the region of tumor sample copy number change (Fig 8B), both Cluster B and Cluster A & C had frequent amplification on chromosome 8q and 18q, and frequent deletion on chromosome 9q. In addition, a significant deletion of Cluster B occurred on chromosome 18q, suggesting that the deletion on chromosome 18q may be associated with poor prognosis. Many studies using different molecular techniques have agreed that the loss of chromosome 18Q is an early event in pancreatic cancer. Chromosome 18Q contains a group of tumor or metastasis-suppressor genes such as SMAD2, SMAD4, DCC, and PAI-2. Loss of 18q heterozygosity (LOH) is a common event in over 90% of pancreatic cancers, while only 50% of pancreatic cancers are characterized by bi-allele inactivation of the SMAD4 gene [39, 40]. In pancreatic cancer, KRAS, TP53, p16, and SMAD4 are thought to play key roles in tumorigenesis [41]. As with many aspects of PDAC, mutated KRAS is to blame. KRAS is an early oncogenic event that drives the recruitment of B cells early in PDAC development and thus plays an immunosuppressive role [42]. Therefore we thought the frequent deletion of chromosome 18q may influence KRAS, TP53, p16, and SMAD4 and induce worse prognosis. KRAS, TP53, CDKN2A, and SMAD4 are the most commonly altered genes in human PDAC, but are considered to have a poor therapeutic response in most cases. , no targeted therapy for PDAC has yet been approved, and despite the large investment in the development of KRAS pathway inhibitors, trial results to date have been unsatisfactory [43]. The counts of Cluster C were much lower than the other two clusters, which may influence the immune therapy effect. We should collect more data to confirm the real therapy effect.

We conclude that Cluster B had a worse prognosis than others groups. Then we designed a classification model to testify whether the sample is Cluster B. In Cluster B, high infiltrating resting NK cells in the tumor microenvironment induce a worse prognosis. We thought the frequent deletion of chromosome 18q may influence KRAS, TP53, p16, and SMAD4 to induce worse prognosis.

In the modern era of immunotherapy, several clinical trials using immune checkpoint inhibitors have been conducted in PAAD and are planned in combination with immunogenic chemotherapy or radiation therapy. Our study shows that taking into account both the type of immune cells infiltrating the tumor and the dominant state of the tumor will lead to more accurate treatment decisions and improved responses to these new therapeutic strategies.

Abbreviations

PAAD (Pancreatic adenocarcinoma)

TME (Tumor microenvironment)

TCGA (The Cancer Genome Atlas)

UCSC (California Santa Cruz)

GSEA (Gene Set Enrichment Analysis)

PDAC (Pancreatic ductal adenocarcinoma)

ECM (Extracellular matrix)

CNV (Copy number variation)

GDAC (Genome Data Analysis Center)

PACA (Pancreatic cancer)

TILs (Tumor-infiltrating immune cells)

STS (Sequence Tagged S)

TIDE (Tumor Immune Dysfunction and Exclusion)

Th1 (T helper 1)

LOH (Loss of heterozygosity)

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable.

Availability of supporting data and materials

Not applicable.

Competing interests

We have no competing interests to declare.

Funding

This study was supported by Special Funding from Guangdong Provincial People's Hospital (2020bq09, 8200100290), Collaborative Funding of Basic and Applied Basis Research Funding of Guangdong Province (2020A1515110536), Basic and applied basic research funding of Guangdong Province (2021A1515011473, 2021A1515012441 and 2021A1515011577); the Science and Technology Program of Guangzhou (202102020030, 202102020107); the Science and Technology Program of Huizhou (2019C0602009), the Science and Technology Program of Huizhou Daya Bay (2020030401), Funding for the Construction of Key Specialty in Huizhou (Qi Zhou), National Natural Science Foundation of China (82072635, 82072637 and 81672475).

Authors' contributions

Shujie Wang: Conceptualization, Methodology, Software, Writing - Original Draft. **Zhenchong Li:** Data curation, Formal analysis, Writing- Original draft.

Chunsheng Liu: Methodology , Visualization, Investigation, Data Curation.

Qi Zhou: Supervision.

Zuyi Ma: Software, Validation.

Zixuan Zhou: Resources.

Hongkai Zhuang: Validation.

Shanzhou Huang: Writing- Reviewing and Editing, Formal analysis.

Chanzhao Zhang: Supervision, Writing - Review & Editing.

Baohua Hou: Writing - Review & Editing, Supervision, Validation.

Acknowledgments

Not applicable.

References

- [1]. Bray, F., et al., Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin, 2018. 68(6): p. 394-424.
- [2]. Hidalgo, M., et al., Addressing the challenges of pancreatic cancer: future directions for improving outcomes. Pancreatology, 2015. 15(1): p. 8-18.

- [3]. Rawla, P., T. Sunkara and V. Gaduputi, Epidemiology of Pancreatic Cancer: Global Trends, Etiology and Risk Factors. *World J Oncol*, 2019. 10(1): p. 10-27.
- [4]. Siegel, R.L., K.D. Miller and A. Jemal, Cancer statistics, 2020. *CA Cancer J Clin*, 2020. 70(1): p. 7-30.
- [5]. American gastroenterological association medical position statement: epidemiology, diagnosis, and treatment of pancreatic ductal adenocarcinoma. *Gastroenterology*, 1999. 117(6): p. 1463-84.
- [6]. Mizrahi, J.D., et al., Pancreatic cancer. *Lancet*, 2020. 395(10242): p. 2008-2020.
- [7]. Clark, C.E., et al., Dynamics of the immune reaction to pancreatic cancer from inception to invasion. *Cancer Res*, 2007. 67(19): p. 9518-27.
- [8]. Provenzano, P.P., et al., Enzymatic targeting of the stroma ablates physical barriers to treatment of pancreatic ductal adenocarcinoma. *Cancer Cell*, 2012. 21(3): p. 418-29.
- [9]. Stromnes, I.M., et al., Targeted depletion of an MDSC subset unmasks pancreatic ductal adenocarcinoma to adaptive immunity. *Gut*, 2014. 63(11): p. 1769-81.
- [10]. Stromnes, I.M., et al., T-cell Localization, Activation, and Clonal Expansion in Human Pancreatic Ductal Adenocarcinoma. *Cancer Immunol Res*, 2017. 5(11): p. 978-991.
- [11]. Topalian, S.L., et al., Safety, activity, and immune correlates of anti-PD-1 antibody in cancer. *N Engl J Med*, 2012. 366(26): p. 2443-54.
- [12]. Royal, R.E., et al., Phase 2 trial of single agent Ipilimumab (anti-CTLA-4) for locally advanced or metastatic pancreatic adenocarcinoma. *J Immunother*, 2010. 33(8): p. 828-33.
- [13]. Thomas, D. and P. Radhakrishnan, Tumor-stromal crosstalk in pancreatic cancer and tissue fibrosis. *Mol Cancer*, 2019. 18(1): p. 14.
- [14]. Alderton, G.K., Microenvironment: An exercise in restraint. *Nat Rev Cancer*, 2014. 14(7): p. 449.
- [15]. Feig, C., et al., The pancreas cancer microenvironment. *Clin Cancer Res*, 2012. 18(16): p. 4266-76.
- [16]. Clark, C.E., et al., Dynamics of the immune reaction to pancreatic cancer from inception to invasion. *Cancer Res*, 2007. 67(19): p. 9518-27.
- [17]. Steele, C.W., et al., Exploiting inflammation for therapeutic gain in pancreatic cancer. *Br J Cancer*, 2013. 108(5): p. 997-1003.
- [18]. Puleo, F., et al., Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features. *Gastroenterology*, 2018. 155(6): p. 1999-2013.e3.

- [19]. Burnet, F.M., The concept of immunological surveillance. *Prog Exp Tumor Res*, 1970. 13: p. 1-27.
- [20]. Ostrand-Rosenberg, S., Immune surveillance: a balance between protumor and antitumor immunity. *Curr Opin Genet Dev*, 2008. 18(1): p. 11-8.
- [21]. Lu, Q., et al., Prognostic significance and immune infiltration of microenvironment-related signatures in pancreatic cancer. *Medicine (Baltimore)*, 2021. 100(12): p. e24957.
- [22]. Ino, Y., et al., Immune cell infiltration as an indicator of the immune microenvironment of pancreatic cancer. *Br J Cancer*, 2013. 108(4): p. 914-23.
- [23]. Wang, L., et al., Prognostic alternative splicing signature reveals the landscape of immune infiltration in Pancreatic Cancer. *J Cancer*, 2020. 11(22): p. 6530-6544.
- [24]. Gentles, A.J., et al., The prognostic landscape of genes and infiltrating immune cells across human cancers. Vol. 21. 2015. 938-945.
- [25]. Ali, H.R., et al., Patterns of Immune Infiltration in Breast Cancer and Their Clinical Implications: A Gene-Expression-Based Retrospective Study. *PLoS Med*, 2016. 13(12): p. e1002194.
- [26]. Bense, R.D., et al., Relevance of Tumor-Infiltrating Immune Cell Composition and Functionality for Disease Outcome in Breast Cancer. *J Natl Cancer Inst*, 2017. 109(1).
- [27]. Gentles, A.J., et al., The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med*, 2015. 21(8): p. 938-945.
- [28]. Subramanian, A., et al., Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. *Proc Natl Acad Sci U S A*, 2005. 102(43): p. 15545-50.
- [29]. Malta, T.M., et al., Machine Learning Identifies Stemness Features Associated with Oncogenic Dedifferentiation. *Cell*, 2018. 173(2): p. 338-354.e15.
- [30]. Tan, T.Z., et al., Epithelial-mesenchymal transition spectrum quantification and its efficacy in deciphering survival and drug responses of cancer patients. *EMBO Mol Med*, 2014. 6(10): p. 1279-93.
- [31]. Hanzelmann, S., R. Castelo and J. Guinney, GSVA: gene set variation analysis for microarray and RNA-seq data. *BMC Bioinformatics*, 2013. 14: p. 7.
- [32]. Ladanyi, A., Prognostic and predictive significance of immune cells infiltrating cutaneous melanoma. *Pigment Cell Melanoma Res*, 2015. 28(5): p. 490-500.
- [33]. Wu, Z., et al., The Landscape of Immune Cells Infiltrating in Prostate Cancer. *Front Oncol*, 2020. 10: p. 517637.

- [34]. Martínez-Rodríguez, M., A.K. Thompson and C. Monteagudo, A significant percentage of CD20-positive TILs correlates with poor prognosis in patients with primary cutaneous malignant melanoma. *Histopathology*, 2014. 65(5): p. 726-8.
- [35]. Moretta, A., et al., Early liaisons between cells of the innate immune system in inflamed peripheral tissues. *Trends in Immunology*, 2005. 26(12): p. 668-675.
- [36]. Martín-Fontecha, A., et al., Induced recruitment of NK cells to lymph nodes provides IFN- γ for TH1 priming. *Nature Immunology*, 2004. 5(12): p. 1260-1265.
- [37]. Ruggeri, L., et al., Effectiveness of donor natural killer cell alloreactivity in mismatched hematopoietic transplants. *Science*, 2002. 295(5562): p. 2097-100.
- [38]. Moretta, L., et al., Killer Ig-like receptor-mediated control of natural killer cell alloreactivity in haploidentical hematopoietic stem cell transplantation. *Blood*, 2011. 117(3): p. 764-71.
- [39]. Hahn, S.A., et al., Homozygous deletion map at 18q21.1 in pancreatic cancer. *Cancer Res*, 1996. 56(3): p. 490-4.
- [40]. Hahn, S.A., et al., DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. *Science*, 1996. 271(5247): p. 350-3.
- [41]. Sunamura, M., et al., Gene therapy for pancreatic cancer based on genetic characterization of the disease. *J Hepatobiliary Pancreat Surg*, 2002. 9(1): p. 32-8.
- [42]. Pylayeva-Gupta, Y., et al., IL35-Producing B Cells Promote the Development of Pancreatic Neoplasia. *Cancer Discov*, 2016. 6(3): p. 247-55.
- [43]. Waters, A.M. and Der CJ, KRAS: The Critical Driver and Therapeutic Target for Pancreatic Cancer. *Cold Spring Harb Perspect Med*, 2018. 8(9).

Tables

Table1. Database of pancreatic adenocarcinoma

Type	Database	Data set	Sample
mRNA	TCGA	PAAD	182
SNV	TCGA	PAAD	178
CNV	Broad Institute GDAC Firehose	PAAD	184
mRNA	ICGC	PACA-CA	234

UCSC, California Santa Cruz; TCGA, The Cancer Genome Atlas; CNV, copy number variation; GDAC, Genome Data Analysis Center; PAAD, Pancreatic adenocarcinoma; PACA, pancreatic cancer.

Table2. Differentially expressed genes in Cluster B when compared to Cluster A & C.

Gene	logFC	adj.P	Gene	logFC	adj.P	Gene	logFC	adj.P
BTK	1.1873	<0.0001	SLAMF6	1.2569	<0.0001	CXCL12	1.1987	<0.0001
PTPRC	1.7012	<0.0001	EBI3	1.0873	<0.0001	MUC1	1.4807	<0.0001
BCL2	1.1112	<0.0001	CD3E	1.4025	<0.0001	CD79B	1.4340	<0.0001
CD5	1.3782	<0.0001	KLRB1	1.2696	<0.0001	MST1R	1.1342	<0.0001
CD1D	1.1403	<0.0001	C7	2.0665	<0.0001	CD79A	1.7005	<0.0001
ITGAL	1.3710	<0.0001	CD2	1.3365	<0.0001	ABCB1	1.1800	<0.0001
CD48	1.5660	<0.0001	SELL	1.4547	<0.0001	CCL17	1.3768	<0.0001
CSF2RB	1.2611	<0.0001	CCL19	2.3173	<0.0001	CXCL9	1.3238	<0.0001
IL7R	1.6148	<0.0001	CCL22	1.1905	<0.0001	CCL21	1.5189	<0.0001
CD1C	1.4496	<0.0001	CD36	1.2128	<0.0001	MIF	1.0984	<0.0001
CCR5	1.1625	<0.0001	CD3D	1.2406	<0.0001	LCN2	1.0433	0.0001
CD84	1.0469	<0.0001	FCER1A	1.2243	<0.0001	CXCL13	1.3897	0.0002
CD8A	1.3678	<0.0001	CD200	1.1305	<0.0001	IL22RA2	1.0967	0.0076
CCL4	1.1943	<0.0001	ITGB4	1.2909	<0.0001	DMBT1	1.0567	0.0061
IL10RA	1.1088	<0.0001	CYFIP2	1.0157	<0.0001	CR2	1.2247	0.0090
IL16	1.1007	<0.0001	RUNX3	1.0235	<0.0001	C8B	1.5060	0.0186
CCR7	1.7241	<0.0001	CD37	1.0938	<0.0001	SPANXB1	3.5421	0.0170
CD27	1.4692	<0.0001	S100B	1.1960	<0.0001			
GZMK	1.4476	<0.0001	MRC1	1.3508	<0.0001			

FC, fold change; adj., adjust.

Figures

Figure 1

Clusters of immune genes in PDAC from TCGA-PAAD. (A) Silhouette plot analysis showed that 3 clusters were the best segmentation of the nCounter. (B) In all TCGC samples were clustered by the Hierarchical clustering method. (C) Unsupervised clustering using correlation distance and ward. D linkage of the correlation matrix assesses the relation between patients according to the expression of the genes on the PanCancer Immune Profiling array. All 758 genes on the array were used for clustering the cohort, annotations of the samples on top of the heatmap indicate histopathological features: pathologic_M and cancer stage as well as the three clusters identified by the cutree method.

Figure 2

Comparison of stromal scores, immune score, estimated scores, and tumor purity between the three clusters. (A-D) Comparison of stemness scores, stromal score, immune score, and estimated scores between the three subtypes (Mann–Whitney U test). P values are from T tests. (E, F) Comparison of survival prognosis between the three subtypes (log-rank test). (G) Spearman correlation coefficients for the two immune subgroups and the prognosis scores in TCGC database. (H) Spearman correlation coefficients for smoking patients in two clusters. (I) Different tumor sites in the pancreas correlated the prognosis scores in the two clusters.

Figure 3

Verification of the model via ICGC data. (A) Spearman correlation coefficients for the two immune subgroups and the prognosis scores in ICGC database. (B) Spearman correlation coefficients for smoking patients or nonsmoking patients correlated the prognosis scores in the two clusters. (C) Different tumor sites in the pancreas correlated the prognosis scores in the two clusters. (D-G) Spearman correlation coefficients for prognosis scores with metastatic tumor, tumor stage, tumor dimension patients with or without diabetes.

Figure 4

Immune clusters are associated with prognosis. (A-C) Kaplan–Meier survival curves for Cluster B (light blue), Cluster A (blue), Cluster C (red), and Clusters A and C (purple). Each Cluster compared with in ICGC database. (D) Cluster B compared with Cluster A & C in ICGC database. Using binomial logistic regression penalized by the lasso method, we trained in ICGC database to predict Cluster B. ROC curve assesses how the lasso output discriminates a sample to be Cluster B or not. p values are from log-rank test.

Figure 5

Immune cell clusters. (A) Density plot represent the distribution of the absolute CIBERSORT scores for selected cell types across the clusters for the METABRIC cohort; the vertical lines crossing the distribution identify the median value for the score. Kruskal–Wallis test p values are denoted. (B) Comparison of immune cell infiltration in two clusters (ANOVA test). (C) The heatmap obtained allows to visualize which cell types are enriched across the immune clusters. Immune clusters are annotated on the top and bottom of the heatmap* $P < 0.05$, ** $P < 0.005$, *** $P < 0.001$.

Figure 7

Gene mutations. (A) The 20 highest mutation genes on the array were used for clustering the cohort, annotations of the samples on the top of the heatmap indicate histopathological features: Age, Race, PN, PM, PT, tumor grade, and tumor stage as well as the three clusters identified by the cutree method. (B, C) Copy number distribution in all tumor samples in two clusters.

Figure 8

TIDE scores. (A) The interwaved violin plot showed that Cluster A & C had little correlation with Cluster B. (B) Cluster C was related to Cluster A and Cluster B. The p value was indicated in detail.

Supplementary Files

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

- [SupplementaryTable1.txt](#)
- [SupplementaryTable10.txt](#)
- [SupplementaryTable2.txt](#)
- [SupplementaryTable3.txt](#)
- [SupplementaryTable4.txt](#)
- [SupplementaryTable5.txt](#)
- [SupplementaryTable6.txt](#)
- [SupplementaryTable7.txt](#)

- [SupplementaryTable8.txt](#)
- [SupplementaryTable9.txt](#)