Available Online: 8 April 2026 DOI: 10.54254/3049-5458/2026.32580

T cell genetic profile changes in pancreatic cancer

Nan Wang

Independent Researcher, East Brunswick, USA

nw2153@nyu.edu

Abstract. This study utilizes single-cell RNA sequencing (scRNA-seq) to examine the immune landscape of Pancreatic Ductal Adenocarcinoma (PDA), adjacent healthy tissues, and Peripheral Blood Mononuclear Cells (PBMCs), focusing on CD4+ T cells, CD8+ T cells, and other T cell subsets in tumor and healthy samples. Using Seurat for data processing and Monocle3 for trajectory analysis, identified distinct immune cell populations and disease-related pathways were identified. UMAP and PCA revealed separation between tumor and healthy tissues, suggesting immune evasion in the tumor microenvironment. Monocle3 analysis showed complex branching in tumor-infiltrating T cells, indicative of activation and immunosuppressive states. Differential gene expression highlighted key immune-regulatory pathways upregulated in the tumor, supporting immune modulation within the tumor microenvironment. Despite the small sample size of healthy cells, this study provides valuable insights into immune interactions in PDA and PBMC. Future research should focus on larger cohorts to validate these findings and explore therapeutic strategies targeting immune cells in the tumor microenvironment.

Keywords: single-cell RNA sequencing, pancreatic cancer, T cell dysfunction, tumor microenvironment, differential gene expression

1. Introduction

Pancreatic cancer presents a significant global health challenge due to its highly aggressive nature and dismal prognosis [1, 2]. This malignancy is characterized by rapid tumor progression, a strong propensity for early metastasis, the frequent absence of noticeable symptoms during its early stages and resistance to conventional therapies [3-5]. All those factors contribute to the challenges in diagnosing and managing the disease. Most patients are diagnosed at advanced stages when curative treatment options are limited. Consequently, these make pancreatic cancer one of the most lethal cancers, and the survival rate for pancreatic cancer remains among the lowest for all cancers, with a 5-year survival rate of less than 10% in many populations [6-9].

Currently advanced therapeutic strategies for pancreatic cancer include surgical resection, radiation therapy, chemotherapy, and more recently, targeted molecular therapies, and immunotherapy [10-12]. While these treatments can provide benefit in certain cases, challenges such as the anatomical location of the pancreas and the asymptomatic nature of early-stage disease continue to hinder early detection efforts [13]. Their overall efficacy is significantly hampered by late-stage diagnoses, the intrinsic resistance of pancreatic cancer to conventional therapies, and the dense desmoplastic stroma that limits drug delivery [14, 15].

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

Immunotherapy has emerged as a promising approach for treating various cancers, including pancreatic cancer. While immune checkpoint inhibitors, such as anti-PD-1 and anti-PD-L1 antibodies, have demonstrated significant efficacy in other malignancies, their effectiveness in pancreatic cancer remains limited due to the tumor's robust immune evasion mechanisms [16, 17]. To resolve this difficulty, and optimize tumor therapeutic outcomes, combination therapy is urgent request. Incorporating immune checkpoint inhibitors, such as CTLA-4, TIM-3 and LAG-3, with chemotherapy, radiation therapy or other traditional therapy methods [18]. As far as novel immunotherapies like Chimeric Antigen Receptor T (CAR-T) cell therapy have shown visionary promise. Although CAR-T therapy has completed remarkable success in hematological malignancies, its harness in solid tumors, including pancreatic cancer, has been restricted by the immunosuppressive Tumor Microenvironment (TME) [19]. New approaches aimed at improving CAR-T cell persistence and overcoming TME-induced suppression are currently being explored, enhancing hope for significant advancements in pancreatic cancer treatment [20].

The Tumor Microenvironment (TME) in pancreatic cancer is a highly intricate and evolving dynamic system made up of cancer cells, immune cells, stromal cells (including tumor-associated fibroblasts, TAFs), endothelial cells, Extracellular Matrix (ECM) components, and a range of secreted substances like cytokines and growth factors [21-23]. These complex compositions interact with each other to create a pro-tumorigenic environment which is beneficial for tumor growth, metastasis, and promote immune evasion, even contributing to resistance to many therapy methods [24, 25].

A significant characteristic of pancreatic cancer TME is the immunosuppressive nature, that could impair the function of immune cells, such as T cells. T cells are crucial in identifying and eliminating disease-associated cells. However, within the pancreatic TME, T cells frequently undergo functional exhaustion due to inhibitory signaling pathways, prolonged antigen exposure, and the presence of immunosuppressive cytokines [26]. Immune checkpoint molecules, such as PD-L1, expressed by tumor cells can interact with PD-1 receptors on T cells, delivering inhibitory signals that hinder T cell activation and proliferation [27]. This interaction leads to T cell dysfunction, lowering their cytotoxic activity and weakening their capacity to generate a robust immune response against the tumor. Complexity of the TME of pancreatic cancer makes the disease much more difficult to conquer.

Due to the complexity of this disease, novel diagnostic tools are urgently needed to enhance early detection and therapeutic outcomes, eventually increasing survival rates [28]. Especially, early diagnosis has the benefit of facilitating timely therapy interventions, broadening the treatment period window, providing more therapeutic potions, and significantly enhance the outcomes. Overcoming these obstacles needs interdisciplinary research, technological innovation and research backup. These improvements are crucial for minimizing the mortality rate related to pancreatic cancer.

T cells are important in the immune system, so that they can recognize and eliminate cancer cells [29]. In normal condition, when T cells recognize Tumor-Associated Antigens (TAAs) presented on the surface of tumor cell by Major Histocompatibility Complex (MHC) molecules, T cells will be active. Once activated, T cells start to play their functions, proliferating and differentiating into effector cells. Effector cells can directly kill ill cells through mechanisms like apoptosis and secrete cytokines to recruit and activate other immune cells, thus, a coordinated full immune response will be formed [30, 31]. T cells' performance to accurately recognize tumor antigens and sustain their attack plays a crucial role in the determination of the total immune response [32]. In pancreatic cancer, T cells frequently become dysfunctional and exhausted because of prolonged antigen exposure, immunosuppressive cytokines, and inhibitory signals within the tumor microenvironment, which hampers their ability to mount a strong immune response against the cancer.

In pancreatic cancer, T cells become dysfunctional and exhausted due to chronic exposure to antigens, the presence of immunosuppressive cytokines, and inhibitory signals in the tumor microenvironment, which impairs their function to effectively combat the tumor [33]. Exhausted T cells express elevated levels of inhibitory receptors like PD-1, CTLA-4, and TIM-3, which impair their function and facilitate immune evasion by the tumor [34]. This effect is most notable in pancreatic cancer, where continuous stimulation of the tumor microenvironment suppresses an effective immune response. T cells in pancreatic cancer are also immune tolerant and see tumor cells as "self" and not foreign, hence further compromising the immune system to develop an effective anti-tumor response [35]. The T-cell deficiency not only endangers the efficacy of these cells to kill tumor cells but also promotes carcinogenesis, establishing an evil cycle, which defeats immune therapies [36]. Comparisons and contrasts among the T-cell genetic blueprints of healthy individuals and patients suffering from cancer may provide key findings on how and why T-cell dysfunction mechanisms happen [37]. This strategy might reveal novel therapy targets or disclose why the patient is resistant to immunotherapy. By identifying the most significant genetic and molecular changes, methods to restore T-cell function, augment immune responses, and even make immunotherapy cancer treatments more effective can be found [38].

New breakthroughs in genomic technologies, specifically single-cell RNA sequencing (scRNA-seq), offer a rare chance to examine the molecular and cellular features of pancreatic cancer at the single-cell level. This technique has transformed the field of cancer biology by making it possible to thoroughly characterize gene expression profiles and identify crucial genetic and epigenetic changes related to tumor development, growth, and drug resistance [39]. The Tumor Microenvironment (TME) is critical in supporting cancer development and immune escape, and in pancreatic cancer, scRNA-seq can be applied to identify diverse T cell subpopulations, activation state, and the molecular mechanisms for T cell dysfunction [40]. T cells in the TME frequently become functionally exhausted, and scRNA-seq can identify the hallmark gene expression signatures of exhausted T cells.

Introduction of scRNA-seq in the diagnostic process holds tremendous promise for the identification of more sensitive and stable biomarkers that may lead to early diagnosis of pancreatic cancer even before the appearance of major symptoms and thus improve prognosis [41, 42]. scRNA-seq also possesses the potential for the detection of yet unidentified targets and molecular mechanisms involved in pancreatic tumorigenesis and immune escape and thus propel the discovery of new therapeutic targets [43]. For instance, by defining immune cell subpopulations and states with a specific focus on T cell dysfunction, scRNA-seq can guide more effective designs for personalized immunotherapy [44].

Furthermore, integration of scRNA-seq with other omics technologies such as proteomics and genomics can be used to uncover new therapeutic targets and biomarkers that can be used for early detection, treatment strategy, and forecasting of prognosis [45, 46]. Integration of scRNA-seq into clinical practice has the potential to revolutionize precision medicine application in pancreatic cancer towards designing personalized therapeutic regimens targeting the tumor and microenvironment's distinct molecular features [47].

In this study, we aim to leverage scRNA-seq technology to uncover the gene underpinnings underlying T cell dysfunction in pancreatic cancer. Through identifying novel gene targets of T cell exhaustion, we aspire to inform the development of therapies that salvage T cell function, reduce tumor immune evasion, and enhance the effectiveness of current and future therapies. Through these projects, we anticipate a greater understanding of the immune profile of pancreatic cancer and contribution to more effective and personalized treatments for this challenging disease.

2. Results

2.1. Methods

Data analysis and processing were done in R (version 4.0) using the Seurat and Monocle3 packages. Visualization analysis was performed using ggplot2 and dplyr.

Further, cells with more than 5% mitochondrial gene expression were excluded from analysis since high mitochondrial gene expression is generally associated with stressed or damaged cells that are usually dysfunctional or apoptotic. Such cells do not reflect normal biological conditions. Exclusion of these cells improves data quality and interpretability by reducing noise and enhancing the reliability and validity of the results.

The procedure included feature identification of highly variable features, scaling of data, Principal Component analysis (PCA) for dimensionality reduction, and application of FindNeighbors function to quantify cell similarity. The cells were grouped according to these similarities, and Uniform Manifold Approximation and Projection (UMAP) was applied for additional dimensionality reduction for visualization

To integrate the Pancreatic Ductal Adenocarcinoma (PDA) and Peripheral Blood Mononuclear Cell (PBMC) datasets and remove any batch effects, the Harmony algorithm was used. Harmony is a data integration approach that normalizes multi-source datasets to correct for batch effects. After batch effect correction, intercellular distances were re-calculated, and Seurat's clustering algorithm was executed to re-cluster the cells. PCA was then used for dimensionality reduction, and finally UMAP for two-dimensional visualization. A comparison of cell distribution before and after integration was made to assess data integration success.

Cell type annotation was performed using the Azimuth method, which predicts cell types based on a reference dataset. The predicted cell types were used to label cells in the visualization, facilitating the identification of immune cell populations, including T cells, in both tumor and peripheral blood samples. The results showed that cells of the same predicted type were spatially clustered, confirming the reliability of the annotation.

T cells were isolated separately for downstream analysis of T cell gene profile changes. Marker genes distinguishing healthy and cancerous cells were identified using the FindMarkers function. The differences in gene expression between cancer and healthy groups were visualized using heatmaps. However, no statistical filtering was applied at this stage.

Differential Gene Expression (DGE) analysis was performed using DESeq2 to identify Differentially Expressed Genes (DEGs) between cancer and healthy groups. Specifically, DEG analysis was conducted for CD4+ and CD8+ T cells in blood samples, as well as for CD4+ T cells, CD8+ T cells, and other T cell subsets in tissue samples (given the smaller sample size). These T cell subpopulations are particularly relevant in pancreatic cancer research. Genes were considered significantly differentially expressed if they met the threshold of p-value < 0.05 and an absolute log2 fold change > 1. Volcano plots were generated to highlight statistically significant genes, providing an intuitive representation of the results.

For DEGs, Gene Ontology (GO) enrichment analysis was conducted to explore the biological processes associated with these genes. First, gene symbols were converted into ENTREZ IDs using the org.Hs.eg.db database. GO terms related to biological processes were then analyzed to identify enriched pathways among the significant genes. Dot plots and category network plots were used to visualize the top GO terms and their associated genes. Notably, tissue samples revealed enrichment of multiple T cell-related GO terms.

Trajectory analysis of T cells was conducted using Monocle3. Monocle3 objects were first constructed to represent T cell gene expression data, followed by classification of cells along a developmental trajectory.

PCA was used for dimensionality reduction, and UMAP was applied for visualization of T cell clustering. The learn_graph function was then used to compute cell similarities and infer the developmental trajectory of T cells. A pseudotime trajectory was constructed to model the differentiation processes of T cells in both healthy and cancer samples.

2.2. Analysis results

The data for this study were obtained from the NCBI database (National Center for Biotechnology Information) and originate from the dataset titled Multimodal Mapping of the Tumor and Peripheral Blood Immune Landscape in Human Pancreatic Cancer, published on October 1, 2020 [33]. This dataset provides single-cell information from pancreatic cancer patients and healthy individuals, processed using single-cell sequencing on the HiSeq 4000 or NovaSeq 6000 platforms with a sequencing depth of 100,000 reads. It includes 16 Pancreatic Ductal Adenocarcinoma (PDA) tissue samples, 3 adjacent normal pancreatic tissue samples, 16 Peripheral Blood Mononuclear Cell (PBMC) samples from pancreatic cancer patients, and 4 PBMC samples from healthy donors, offering valuable insights into the tumor and peripheral immune microenvironment of pancreatic cancer.

After merging all the data using the Seurat package in R, two datasets were obtained: Pancreatic Ductal Adenocarcinoma (PDA) and Peripheral Blood Mononuclear Cells (PBMC). Each dataset includes both healthy and cancer groups.

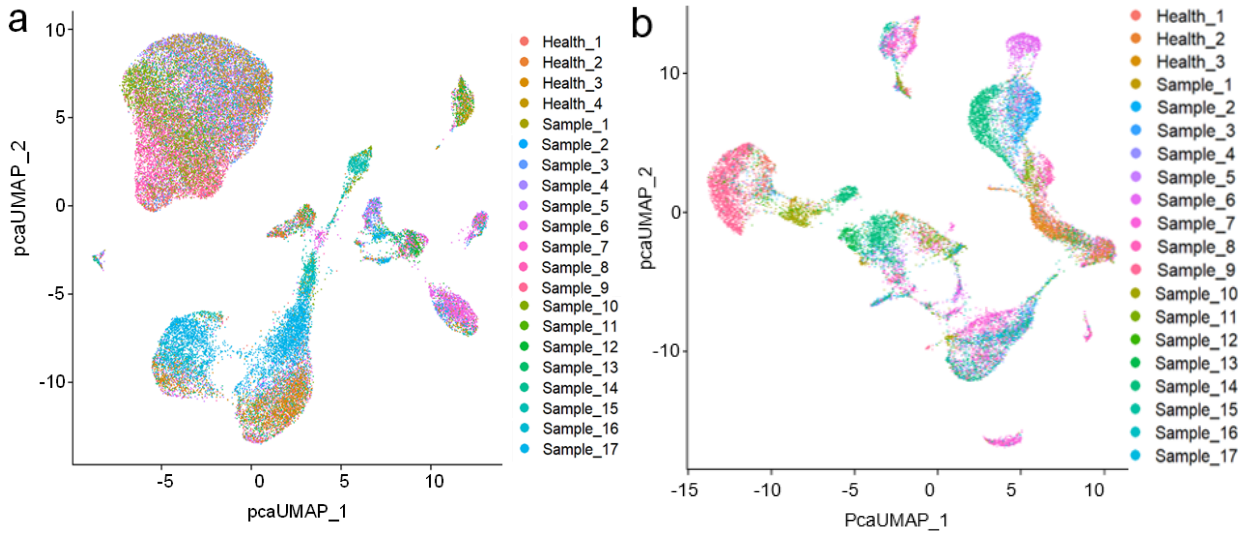

Figure 1. UMAP plots from raw data

UMAP projection of (a) PBMC samples and (b) PDA samples based on data source clustering. Cells are grouped according to their original data sources, with colors representing different sample groups.

Each dataset yielded a UMAP plot showing the distribution of samples across two dimensions. The plots clearly show distinct separation between the samples, suggesting the potential presence of batch effects between the datasets. Especially for tissue sample (Figure 1b), the sample distribution difference is obvious. Although blood sample (Figure 1a) seems present relatively good integration, further butch effects elimination should be done.

Subsequently, the Harmony method was applied to integrate the two datasets to eliminate batch effects. Based on the integrated data, the distances between cells were recalculated, and the cells were re-clustered.

UMAP was then used again for dimensionality reduction of the data.

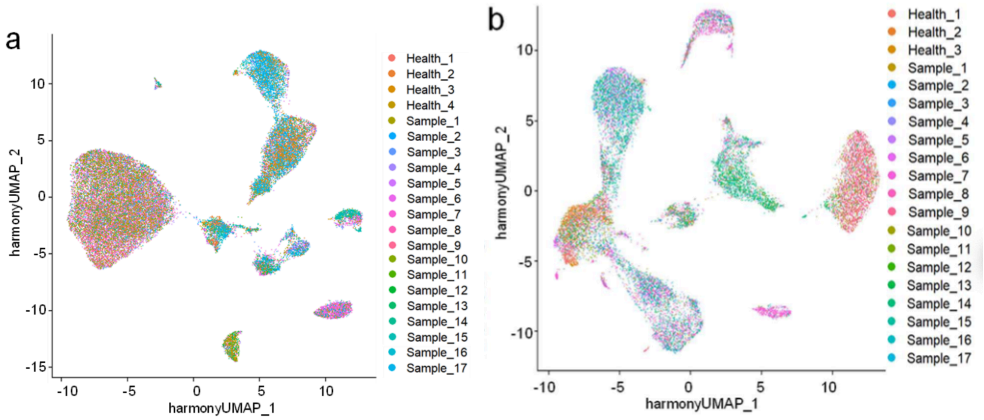

Figure 2. UMAP plots after harmony integration

UMAP projection after harmony integration adjudication of (a) PBMC samples and (b) PDA samples. Colors represent different sample groups.

Two new UMAP plots were generated showing the distribution of cells across different samples but after harmony integration. Comparing these plots with the previous ones, both blood sample (Figure 2a) and tissue sample (Figure 2b), are evident that the cells have achieved better integration. Neighboring cell points now originate from different samples, and different sample sources have been well integrated, indicating that the batch effects have been successfully mitigated.

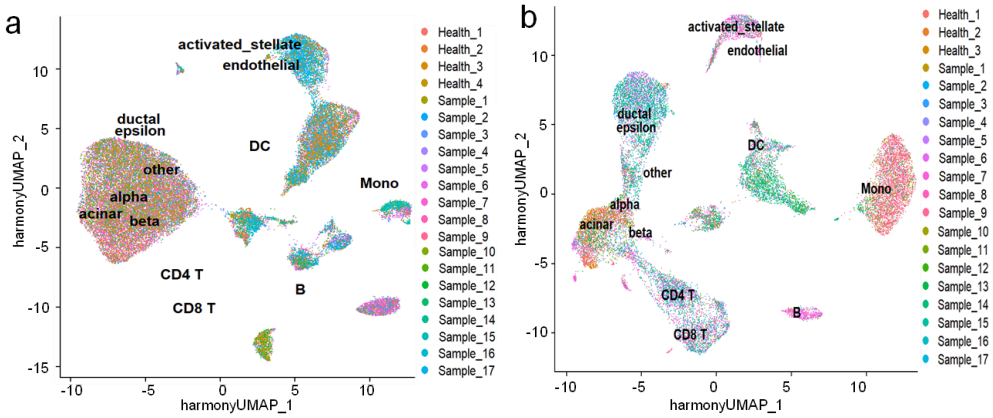

Figure 3. UMAP plots with cell type annotations

UMAP projection of (a) PBMC samples and (b) PDA samples with Azimuth predicted cell type annotation. Colors represent different sample groups.

Using the cell type annotations predicted by AZIMUTH, cell type information was added to the UMAP plots. Blood sample (Figure 3a) only contains immune cells, while tissue sample (Figure 3b) contains other cells. By visualizing the distribution of specific cell types in the UMAP space, it was observed that cells of the same type tend to cluster together. This validated the AZIMUTH annotation method worked well. Furthermore, combining different sample source information (Figure 2a and Figure 2b), it is obvious that cells

of the same type were found to be present in multiple samples, with a uniform distribution across different samples.

As expected, different cell types exhibited relatively distinct distributions, highlighting the gene expression differences between various cell types. This also supports the notion that the integration process has successfully minimized batch-related variations, allowing for a clearer representation of biological variability across cell types.

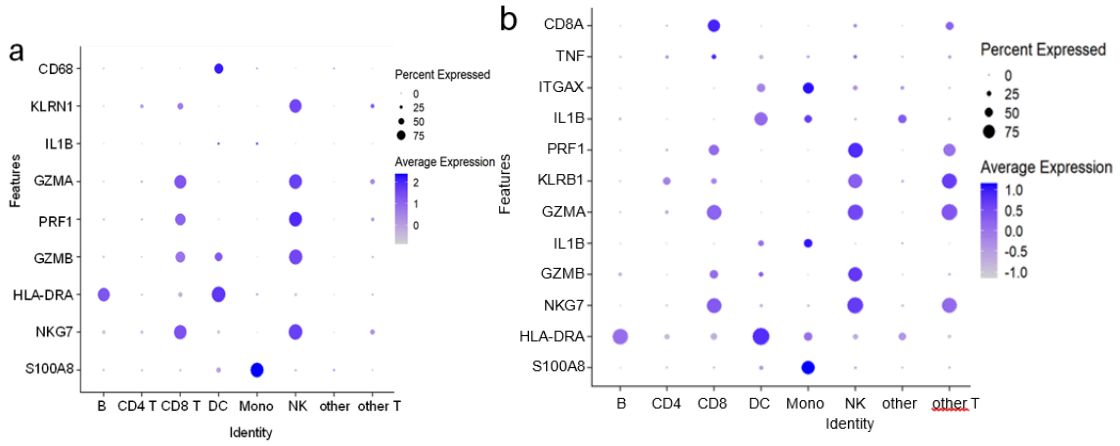

Figure 4. Dot Plot of immune cell markers

The dot plot visualizes the expression levels of selected marker genes across different predicted immune cell types of (a) PBMC samples and (b) PDA samples. The x-axis represents predicted cell types, while the y-axis displays the selected marker genes. Dot size corresponds to the percentage of cells expressing each gene within a given cell type, and color intensity represents the average expression level of each gene.

The dot plot visualizes the differentially expressed genes of immune cells, highlighting key molecular and gene distinctions between different immune cell types. The color intensity represents the average expression level of each gene within a given cell group, with darker shades indicating higher expression, while the dot size corresponds to the proportion of cells expressing the gene, with larger dots representing broader expression across the population. This analysis identifies genes that are both highly expressed and widely distributed within specific cell types, providing insights into potential biomarkers and key regulators of cellular heterogeneity and functional diversity. From blood sample (Figure 4a) and tissue sample (Figure 4b), the expression of the KLRB1 gene is more prominent in CD4 T cells, while the expression of NKG7, GZMB, GZMA, and PRF1 genes is generally more prominent in CD8 T cells. In other T cells, the expression of GZMA, KLRB1, PRF1, and NKG7 genes is relatively significant. These genes may provide clues for identifying the pathways of T cell gene alterations in pancreatic cancer in subsequent analyses.

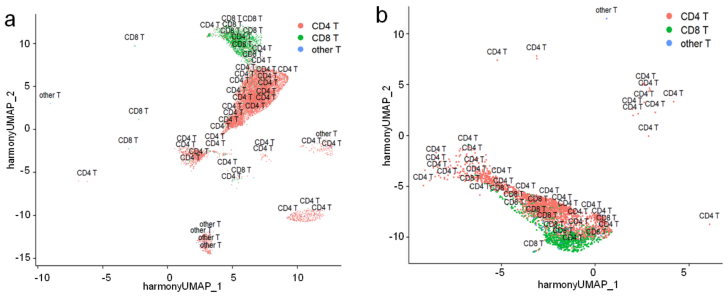

Figure 5. UMAP with T cell type annotations

UMAP projection of T cells from (a) PBMC samples and (b) PDA samples, annotated with Azimuth-predicted cell types. Colors indicate different cell types.

T cells were subset from the total dataset, including CD4, CD8, and other T cell types. For these T cells, clustering was performed based on the distances computed from PCA by Seurat, and a UMAP plot was generated to visualize their distribution in two-dimensional space.

From the plots, it is evident that different t cell clusters form clear separations. At the same time, T cells of the same type tend to aggregate together, indicating that the clustering process is accurate and able to group similar cells into the same cluster.

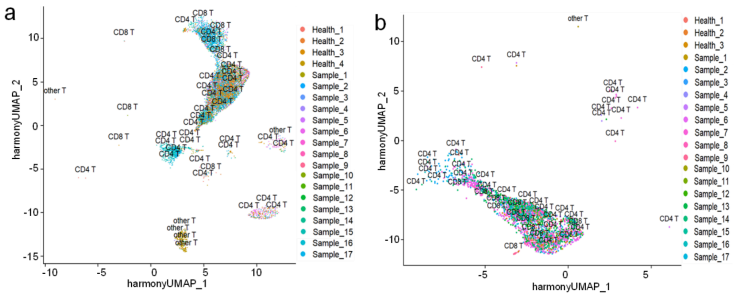

Figure 6. UMAP with T cell type annotations group by sample source

UMAP projection of T cells from (a) PBMC samples and (b) PDA samples, with the annotation of Azimuth-predicted cell types and grouped by data sources.

At this stage, T cells were grouped by sample source, and their distribution in the UMAP two-dimensional space was visualized. Cells from each sample are evenly distributed within the same T cell subtype clusters. This indicates a consistent representation of cell types across samples and reflects the biological distinction between different T cell subgroups. It can also be observed that T cells from the healthy group are less abundant compared to those from the cancer group. This is attributable to the smaller amount of raw data available for the healthy group, which explains the lower representation of T cells in this group. It may also be because the immune system of patients in the cancer group is in a highly active state, which results in a higher number of T cells compared to the healthy group.

Combined with the previous Harmony results (Figure 2), which demonstrated the successful elimination of batch effects, this consistency further supports the presence of significant differences between T cell types. The annotations of cell types in these two plots (Figure 6) validate the alignment between cell clustering and sample distribution, confirming their correspondence with distinct cell types. This alignment underscores the robustness of the clustering and the accuracy of the sample-based grouping in capturing cellular heterogeneity.

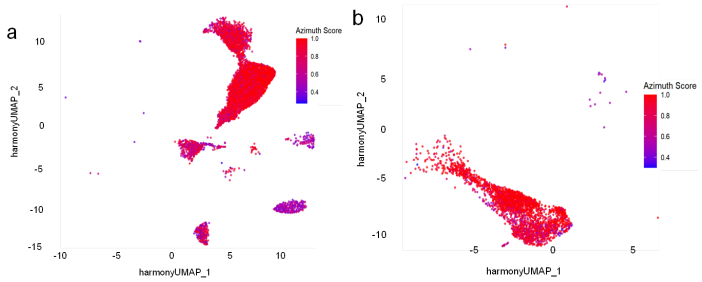

Figure 7. Azimuth score plots

The plots show the UMAP representation of cells from the harmony UMAP reduction, with each point colored according to its Azimuth predicted cell type score. The color gradient ranges from blue (low Azimuth score) to red (high Azimuth score), indicating the strength of the predicted cell type classification. Left plot is for cells from (a) PBMC samples right plot is for cells from (b) PDA samples.

Next, T cell distribution plots were created using ggplot, with the AZIMUTH score represented as a color gradient. The color ranges from blue (low score) to red (high score). The colors in the plot reflect the confidence level of cell type classification, where higher scores indicate more reliable classification results. The majority of the cells in the plots exhibit high AZIMUTH scores, indicating that the classification results are reliable for most cells. Especially for the tissue sample (Figure 7b), only few cells have lower azimuth score, while most of the cells with really high score level. For blood sample (Figure 7a), there are more T cells with lower azimuth socre, however, they still exist with moderate or relatively high confidence.This visualization effectively highlights the robustness of the cell type annotations and their confidence levels across the dataset.

Combining the two plots with the previous T cell type distribution and sample distribution plots (Figure 5 and Figure 6), it can be observed that low-confidence cells (blue) are randomly distributed across various cell types and samples. This suggests that problematic cells and data noise are minimal within this dataset, further confirming the reliability and quality of the data and cell classifications.

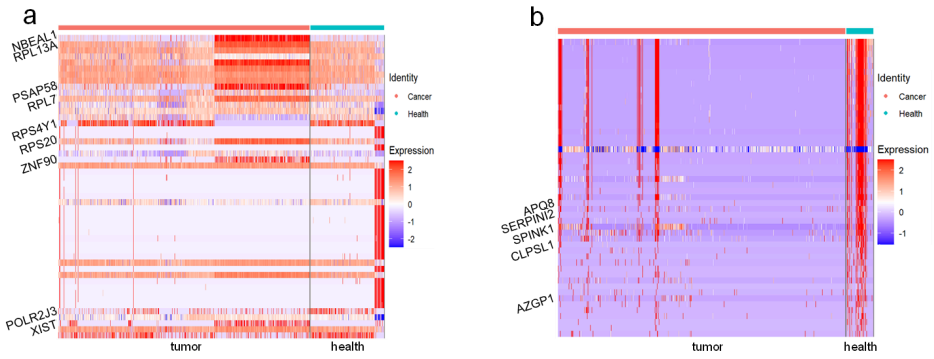

Figure 8. Heatmap for top 50 gene expression

The heatmaps show the expression of the top 50 genes most differentially expressed between the Health and Cancer sample groups for (a) PBMC samples and (b) PDA samples. Each row represents a gene, and each column represents a sample from the two groups (Health and Cancer). The color gradient ranges from blue (low expression) to red (high expression). The plot highlights gene expression patterns across different sample groups, with lines drawn between clusters for better visualization.

Heatmaps were generated to display the expression patterns of the top 50 differentially expressed genes, categorized by cancer and healthy groups. In the heatmap, blue represents low expression levels, white indicates moderate expression, and red signifies high expression. The heatmap reveals notable differences in gene expression between the cancer and healthy groups, which with the gene name labels, such as NBEAL1, XIST for blood sample (Figure 8a) and AZGP1 for tissue sample (Figure 8b). Additionally, genes with high differential expression vary between blood and tissue samples, which could be further researched. Some genes show consistently high expression in cells from the healthy group but are only expressed in a subset of cells from the cancer group, like RPS4Y1 in blood sample (Figure 8a). This observation may highlight distinct biological processes or cellular states associated with the cancer group, shedding light on potential mechanisms underlying disease progression or immune system alterations.

Next, further analysis will focus on the highly differentially expressed genes. All samples were grouped into "healthy" or "tumor" categories, and the average expression of each gene in each group was calculated. Due to the smaller sample size in the healthy group, the analysis primarily focused on CD4 and CD8 T cells combined for blood sample, and CD4, CD8 and other T cells were combined used for tissue sample. Differential Expression (DE) analysis was conducted using the DESeq2 method to identify marker genes for those T cells in the cancer and healthy groups. Genes were filtered based on their *p*-value and average log2 fold-change (avg_log2FC), with significant genes identified using the criteria *p*-value smaller than 0.05 and the absolute value of log 2 fold change larger than 1.

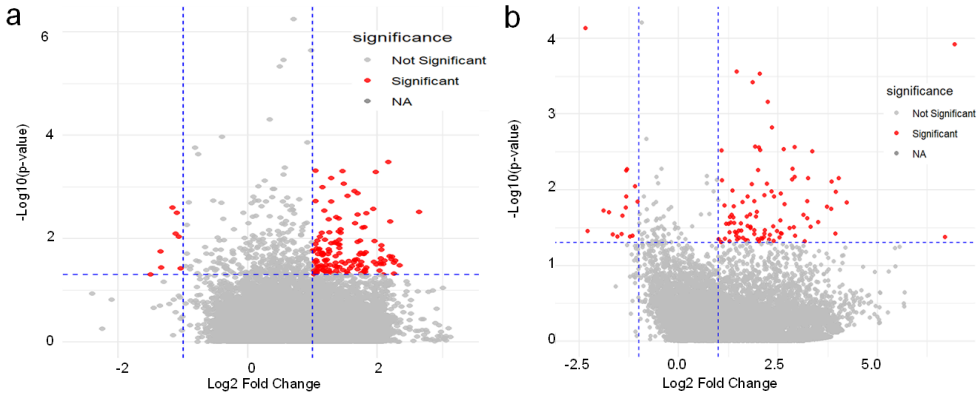

Figure 9. Volcano plots for marker genes

The two volcano plots for (a) PBMC samples and (b) PDA samples. The plots display the log2 fold changes (x-axis) and -log10 *p*-values (y-axis) of differentially expressed genes identified by DESeq2. Genes are colored based on their significance, with red representing genes that are significantly differentially expressed (*p*-value < 0.05 and absolute log2 fold change > 1) and gray representing non-significant genes. Dashed blue lines indicate the significance thresholds for *p*-value (0.05) and log2 fold change (± 1). This plot highlights the relationship between gene expression changes and their statistical significance.

Volcano plots were used to display significant (red) and non-significant (gray) genes, with each point representing a gene. Genes located above the horizontal reference line meet the significance threshold (*p*-value < 0.5), while those outside the vertical reference lines at −1 and 1 indicate substantial expression changes. In the plot, red points are primarily concentrated outside the reference lines, highlighting genes with both statistically significant and biologically meaningful changes. In contrast, non-significant genes are more randomly distributed, mostly within the reference lines, reflecting minimal or insignificant expression

differences. For both of the plots (Figure 9a is for blood sample and Figure 9b is for tissue sample), there are many significant gene expression differences which can be used for further analysis.

Using ENTREZID, gene symbols were converted into identifiers recognizable by the org.Hs.eg.db database. Only genes that were successfully mapped to valid identifiers in org.Hs.eg.db were included in subsequent Gene Ontology (GO) analysis, focusing on their associated biological processes. The resulting analysis provided a comprehensive view of the biological roles these genes play, illustrated in the following image.

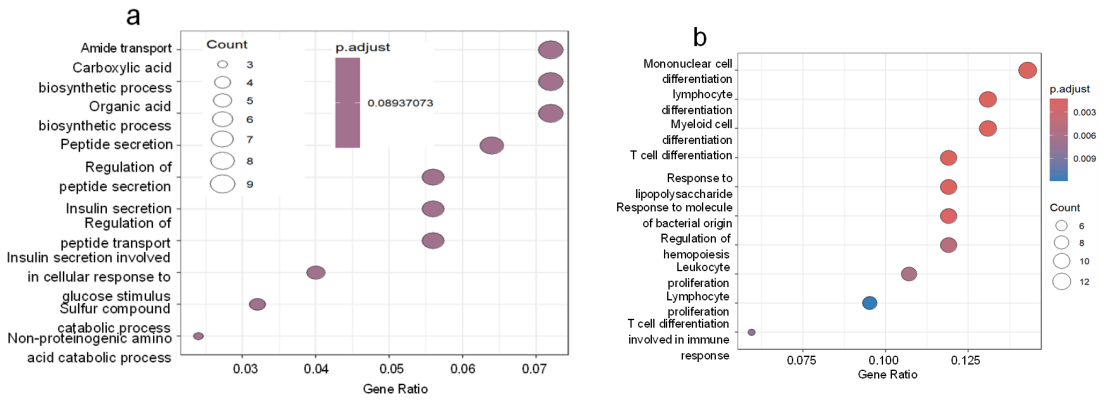

Figure 10. Dot plots for biological processes

The plots display the top 10 enriched Biological Processes (BP) as determined by the GO term enrichment analysis for (a) PBMC samples and (b) PDA samples. The x-axis represents the gene ratio (the proportion of genes annotated to a particular GO term), while the y-axis shows the GO term description. The color scale represents the adjusted *p*-value, with darker colors corresponding to more statistically significant terms (lower *p*-values). The size of the dots represents the number of genes associated with each GO term, with larger dots indicating a higher number of genes.

Dot plots were generated to visualize GO terms and their corresponding enrichment factors, where the enrichment factor represents the ratio of genes associated with a GO term to the total number of genes in the entire set. In the plots, the color of each dot reflects the significance of the GO term, with deeper colors indicating smaller *p*-values (greater significance). The size of each dot represents the number of enriched genes within the GO term, with larger dots corresponding to terms associated with more genes. For blood sample (Figure 10a), all those 10 biological processes have small range of *p*-value, all of those were significant. Additionally, for tissue sample (Figure 10b), the 10 go terms were much more significant, especially for the two: lympho proliferation and T cell differentiation involoved in immune response, which related to the role of T cell proliferation and differentiation in pancreatic cancer. Further analysis can reveal the expression patterns of immune-related genes and identify potential immunotherapy targets for pancreatic cancer. This visualization highlights the most statistically and biologically significant pathways and processes enriched in the dataset.

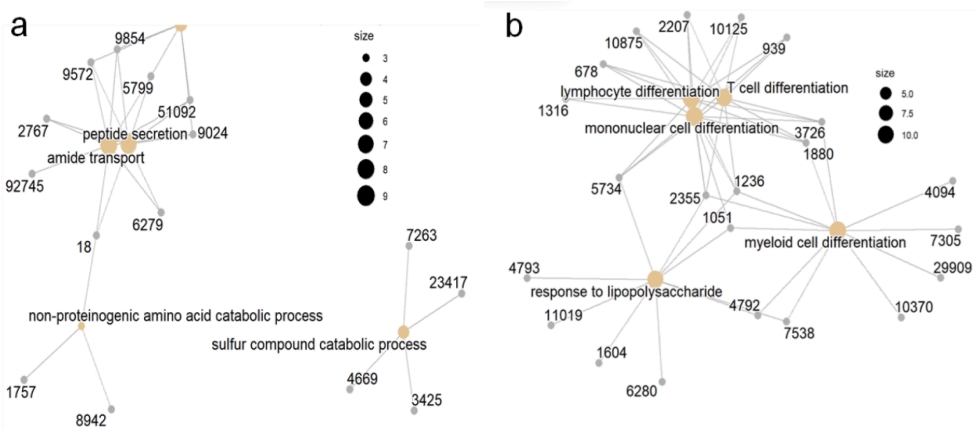

Figure 11. Top 5 enriched GO terms and their associated genes

Gene Ontology enrichment analysis network plots for the top 5 enriched Biological Processes (BP) based on the GO term enrichment analysis for (a) PBMC samples and (b) PDA samples. Each node represents a GO term (biological process), and the size of the node reflects the number of genes associated with that GO term. The edges between the nodes represent gene-term associations, where the genes connected to the same GO term are grouped together.

The cnetplot (circular network plot) visualizes the relationships between genes and GO terms, where top 5 GO terms are represented as a node (circle), and the genes associated with that term are connected to it by edges (lines). The size of the GO term node reflects the number of genes associated with it, with larger nodes indicating terms involving more genes. The genes are displayed as smaller nodes connected to their corresponding GO terms, and the edges show the associations between them. This plot provides an intuitive way to understand how different genes are involved in various enriched GO terms. As for the blood sample (Figure 11a), most of the GO terms were related to amino acids and protein, while, as for tissue sample (Figure 11b), there are terms related to immune response, especially the T cell differentiation.

Next, based on the previously defined T-cell subgroups, Monocle3 objects were created, which includes a gene-by-cell expression matrix, metadata, and gene annotation information. The object was then subjected to PCA for dimensionality reduction, followed by UMAP visualization to explore the data in a lower-dimensional space. Cluster grouping was also performed to identify distinct cellular populations. Using the learn-graph function, trajectories based on the similarity between cells were inferred, resulting in a topological structure that represents the progression or differentiation of T-cells. This analysis revealed insights into the potential cellular trajectories and dynamic changes within the T-cell subgroups, as illustrated in the following images.

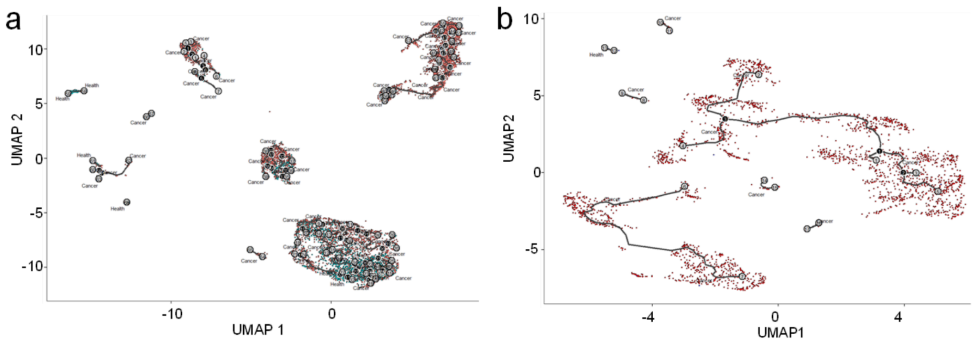

Figure 12. Trajectory plots

UMAP projection plots of trajectory analysis for (a) PBMC samples and (b) PDA samples. For the two graphes, each point represents a single cell, and the plot is colored by the group type (as healthy and tumor). The colors correspond to different groups (red is for tumor cells and blue is for healthy cells), allowing for visualization of the clustering patterns and cell-type distribution. Cells are grouped based on the hierarchical clustering results. Cell clusters were identified using UMAP for dimensionality reduction and a PCA-based preprocessing method.

The analysis reveals paths when cells are colored by cancer and healthy groups. Due to small health sample size, there are only few of healthy cells shown in the plots. Several clusters of cells in the cancer group are distinctly separated, indicating significant differences between these cells. This suggests that cancer cells may have undergone specific biological processes and transitions, such as malignant transformation or immune evasion.

On the far left of blood sample (Figure 12a), a group of cells is predominantly composed of healthy cells, representing the stable state of cells from healthy individuals. However, another group of cells exhibits a mixed distribution, with healthy cells located towards the lower part of the cluster. This indicates a more complex interaction between cancerous and healthy cells, suggesting that the developmental trajectories of cancer and healthy cells may intersect at multiple points. This could imply that cancer cells and healthy cells share certain developmental pathways or undergo overlapping transitions during disease progression.

For the tissue section (Figure 12b), the track was much clearer. Healthy cells were located in the top left corner and clearly separated with tumor cells. There may be extreme differences in gene expression or functional states between cancer cells and healthy cells. This separation may be relevant to changes in tumor development or the immune microenvironment, reflecting the biological differences between cancer cells and healthy cells. Trajectory lines on the graph help us to understand the dynamic change of cell states, revealing the cell differentiation or transformation process at different stages or under different conditions.

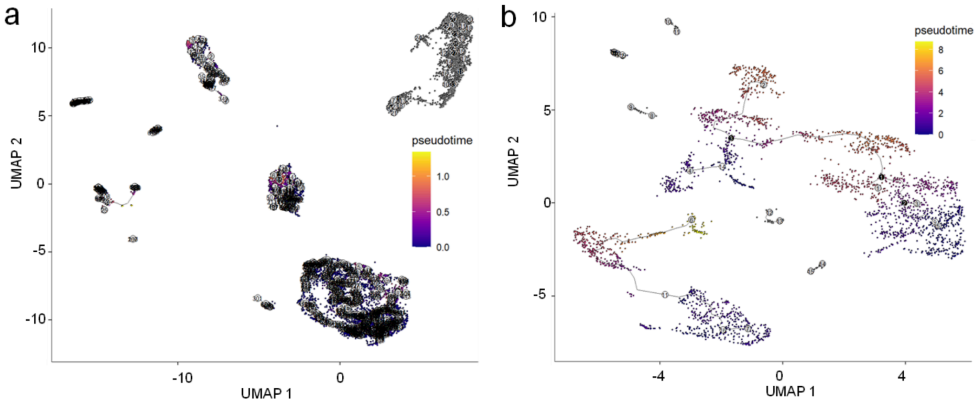

Figure 13. Pseudotime trajectory analysis plots

Pseudotime trajectory analysis for (a) PBMC samples and (b) PDA samples. Cells are ordered along the trajectory based on pseudotime, a measure of the progression of cells along a dynamic biological process. The color scale represents pseudotime, with darker colors indicating cells at earlier stages and lighter colors indicating cells at later stages. The trajectory graph is shown with the root cells defined as "Health" cells.

Based on the pseudotime analysis direction plot, healthy cells were the reference point and sequentially moved to a diseased state along the direction. When pseudotime is progressing, cell color shifts from the steady state of healthy cells to yellow, indicating significant cell state changes with time. This color change not only captures the dynamic cellular biological processes but also reflects the different developmental or

pathological states cell populations experience at different times. In the pseudotime trajectory, cell populations are colored individually along the pseudotime trajectory and can potentially represent different biological stages or have different functions in cell fate specification.

For the blood sample (Figure 13a), some cell populations are closely grouped together, suggesting that these cells may have undergone similar biological processes or transitions. Other cell populations are more dispersed along the trajectory, indicating that they may be at different stages of transition or exhibit more complex heterogeneity during development. This heterogeneity could result from differences in gene expression, changes in the cellular microenvironment, or external stimuli. In particular, in cancer samples, we may observe certain cell populations transitioning from a healthy state to a diseased state, possibly undergoing processes like malignant transformation, genetic mutations, or immune evasion.

To the tissue sample (Figure 13b), cell trajectory pathway is more clearly. The image illustrates that within a large cluster of cells, relatively healthy cells were located in the lower region of the UMAP2 axis. This suggests that these cells represent an early disease state with no significant genetic alterations. In contrast, cells positioned at higher UMAP values exhibit more pronounced changes, as indicated by their shift toward a yellow color gradient. These regions contain a greater proportion of transformed or malignant cells, implying underlying genetic expression changes, phenotypic alterations, or malignant progression.

In addition, the plot of trajectory demonstrates possible intersections and overlapping areas among cancer and normal cells, which indicates that healthy cells and cancer cells could have similar developmental pathways or cell fates in essential biological processes. These intersections might indicate interactions or co-evolutionary relationships between normal and cancer cells at specific stages of transition, which is important for delineating the mechanism of cancer initiation and progression.

3. Conclusion

In this study, single-cell RNA sequencing (scRNA-seq) was used to profile the immune profile of Pancreatic Ductal Adenocarcinoma (PDA), and Peripheral Blood Mononuclear Cells (PBMCs) with a focus on heterogeneity in immune cell composition—CD4+ T cells, CD8+ T cells, and other T cell subsets—within tumor tissues, adjacent normal tissues. The data were processed carefully, normalized, and analyzed computationally to derive insight into the T cell genetic profile changes in PDA and PBMCs samples. We used various computational tools, including Seurat for data preprocessing and clustering and Monocle3 for the analysis of trajectories, to uncover the dynamic changes and developmental pathways of immune cells.

One of the most important findings in this study is the differential immune cell subpopulations in tumor versus normal tissues, most strikingly in the subset of T cells. UMAP and PCA plots demonstrated clear discrimination between tumor and normal tissue-derived immune cells, consistent with the early notion of immune evasion in cancer. Such clear-cut clustering suggests that tumor-infiltrating immune cells undergo separate biological events, such as malignant transformation and immune escape, with distinct gene expression profiles and defective immune interactions. In contrast, immune cells from regular tissues clumped in tighter groups, which indicated a more functional and intact immune microenvironment.

Further trajectory analysis using Monocle3 revealed significant developmental differences between tumor and control T cells. Interestingly, certain tumor-derived immune cell clusters exhibited clear patterns of branching, suggesting the existence of novel cellular states. This implies that during disease progression, T cells might get activated, differentiate, or become immunosuppressive—a common event in the tumor microenvironment. Healthy immune cells, on the other hand, demonstrated a more organized distribution and lower trajectory complexity characteristic of a homeostatic immune environment.

Among intriguing observations in the present study is the partial intermingling of healthy tumor immune cells in certain groups, suggesting the heavy influence of tumor microenvironment on immune cell states. Cell-to-cell contact between tumor cells and the immune cells can be crucially significant in invoking immune suppression at the tumor site, an aspect which holds the key to cancer proliferation. Understanding such engagements may lead to novel concepts for therapeutic intervention aimed at restoring the immune response.

Differential Gene Expression (DEG) analysis also highlighted key molecular pathways distinguishing tumor-infiltrating and normal immune cells. T cells in the tumor microenvironment were enriched with activation and differentiation genes, indicating an immune response state. However, due to immune evasion mechanisms, the functional potential of these T cells may be impaired. Additionally, inflammation and immune regulation-related genes were overexpressed in tumor tissues, corroborating the hypothesis of immune modulation in the tumor microenvironment.

One limitation of this study is the imbalance in cell numbers in the healthy and tumor groups. After stringent filtering of genes, the number of healthy immune cells was very small, and this could introduce bias in differential expression analysis and pseudotime inference accuracy, respectively. Pseudotime analysis could be hampered by a small sample size in the healthy group and thus fail to capture critical stages of development or patterns of differentiation. To address this challenge, future research must have greater healthy cells available for analysis or utilize weighted statistical models to negate group differences.

In conclusion, this study introduces new findings concerning the immune landscape of PDA, and PBMCs, and the complex interactions of immune cells with tumor cells. The integration of single-cell RNA sequencing with computational tools facilitated the identification of hallmark immune cell subsets, their developmental trajectories, and the molecular mechanisms that supported differences between tumor and normal tissues. Subsequent studies will be required to find candidates for biomarkers of immunotherapeutic strategies and validate them in larger patient cohorts. Increased examination of immune cell subpopulations in the tumor microenvironment may provide an avenue for new therapies with significant clinical implications.

References

[1] Ilic M, Ilic I (2016). Epidemiology of pancreatic cancer. *World J Gastroenterol*, *22*(44), 9694-9705. doi: 10.3748/wjg.v22.i44.9694

[2] Zhu X, Niu X, Ge C (2019). Inhibition of LINC00994 represses malignant behaviors of pancreatic cancer cells: interacting with miR-765-3p/RUNX2 axis. *Cancer Biol Ther*, *20*(6), 799-811. doi: 10.1080/15384047.2018.1564566

[3] McGuigan A, Kelly P, Turkington RC, Jones C, Coleman HG, McCain RS (2018). Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes. *World J Gastroenterol*, *24*(43), 4846-4861. doi: 10.3748/wjg.v24.i43.4846

[4] Rouanet M, Lebrin M, Gross F, Bournet B, Cordelier P, Buscail L (2017). Gene therapy for pancreatic cancer: specificity, issues and hopes. *Int J Mol Sci*, *18*(6), 1231. doi: 10.3390/ijms18061231

[5] Kumar R, Patiyal S, Kumar V, Nagpal G, Raghava GPS (2019). In silico analysis of gene expression change associated with copy number of enhancers in pancreatic adenocarcinoma. *Int J Mol Sci*, *20*(14), 3582. doi: 10.3390/ijms20143582

[6] Klein AP (2021). Pancreatic cancer epidemiology: understanding the role of lifestyle and inherited risk factors. *Nat Rev Gastroenterol Hepatol*, *18*(7), 493-502. doi: 10.1038/s41575-021-00457-x

[7] Gunderson AJ, Kaneda MM, Tsujikawa T, Nguyen AV, Affara NI, Ruffell B, Gorjestani S, Liudahl SM, Truitt M, Olson P, Werb Z, Coussens LM (2016). Bruton tyrosine kinase-dependent immune cell cross-talk drives pancreas cancer. *Cancer Discov*, *6*(3), 270-285. doi: 10.1158/2159-8290.CD-15-0827

[8] Xu D, Wang Y, Chen Y, Zheng J (2023). Identification of the molecular subtype and prognostic characteristics of pancreatic cancer based on CD8+ T cell-related genes. *Cancer Immunol Immunother*, *72*(3), 647-664. doi: 10.1007/s00262-022-03269-3

[9] Geng Y, Guan R, Hong W, Huang B, Liu P, Guo X, Hu S, Yu S, Xu H, Liu C, Li H, Zhou Y, Liu X (2020). Identification of m6A-related genes and m6A RNA methylation regulators in pancreatic cancer and their association with survival. *Ann Transl Med*, *8*(6), 387. doi: 10.21037/atm.2020.03.98

[10] Ansari D, Tingstedt B, Andersson B, Holmquist F, Sturesson C, Williamsson C, Andersson R (2016). Pancreatic cancer: yesterday, today and tomorrow. *Future Oncol*, *12*(16), 1929-1946. doi: 10.2217/fon-2016-0010

[11] Pishesha N, Harmand TJ, Ploegh HL (2022). A guide to antigen processing and presentation. *Nat Rev Immunol*, *22*(12), 751-764. doi: 10.1038/s41577-022-00707-2

[12] Yin L, Lu Y, Cao C, Lu Z, Wei J, Zhu X, Chen X, Li J, Wang Y, Zhang Y, Liu H, Sun S, Zhao G (2022). CA9-related acidic microenvironment mediates CD8+ T cell-related immunosuppression in pancreatic cancer. *Front Oncol*, *11*, 832315. doi: 10.3389/fonc.2021.832315

[13] Archange C, Nowak J, Garcia S, Moutardier V, Calvo EL, Dagorn JC, Iovanna JL (2008). The WSB1 gene is involved in pancreatic cancer progression. *PLoS One*, *3*(6), e2475. doi: 10.1371/journal.pone.0002475

[14] Sherman MH, Beatty GL (2023). Tumor microenvironment in pancreatic cancer pathogenesis and therapeutic resistance. *Annu Rev Pathol Mech Dis*, *18*(1), 123-148. doi: 10.1146/annurev-pathmechdis-031621-024600

[15] Vincent A, Herman J, Schulick R, Hruban RH, Goggins M (2011). Pancreatic cancer. *Lancet*, *378*(9791), 607-620. doi: 10.1016/S0140-6736(10)62307-0

[16] Globig AM, Zhao S, Roginsky J, Maltez VI, Guiza J, Avina-Ochoa N, Weizman OE, Buechler MB, Canner D, Ralainirina N, Brown FD, Herold KC, Palm NW (2023). The β1-adrenergic receptor links sympathetic nerves to T cell exhaustion. *Nature*, *622*(7982), 383-392. doi: 10.1038/s41586-023-06568-6

[17] Wu Y, Zhang C, Liu X, He Z, Shan B, Zeng Q, Zhou Y, Zhang W, Li Y, Wang H, Li H, Zhang J, Chen X (2021). ARIH1 signaling promotes anti-tumor immunity by targeting PD-L1 for proteasomal degradation. *Nat Commun*, *12*(1), 2346. doi: 10.1038/s41467-021-22467-8

[18] Malsy M, Graf B, Almstedt K (2019). The active role of the transcription factor Sp1 in NFATc2-mediated gene regulation in pancreatic cancer. *BMC Biochem*, *20*(1), 2. doi: 10.1186/s12858-019-0105-4

[19] Qi C, Xie T, Zhou J, Wang X, Gong J, Zhang X, Chen Y, Wang L, Li Z, Zhao H, Xu J, Zhang Y, Wang J (2023). CT041 CAR T cell therapy for Claudin18.2-positive metastatic pancreatic cancer. *J Hematol Oncol*, *16*(1), 102. doi: 10.1186/s13045-023-01491-9

[20] DeSelm CJ, Tano ZE, Varghese AM, Adusumilli PS (2017). CAR T-cell therapy for pancreatic cancer. *J Surg Oncol*, *116*(1), 63-74. doi: 10.1002/jso.24627

[21] Huang H, Wang Z, Zhang Y, Pradhan RN, Ganguly D, Chandra R, Murimwa G, Wright M, Guha S, Guha P (2022). Mesothelial cell-derived antigen-presenting cancer-associated fibroblasts induce expansion of regulatory T cells in pancreatic cancer. *Cancer Cell*, *40*(6), 656-673.e7. doi: 10.1016/j.ccell.2022.04.011

[22] Miao L, Liu Q, Lin CM, Luo C, Wang Y, Liu L, Yin W, Hu S, Kim WY, Huang L (2017). Targeting tumor-associated fibroblasts for therapeutic delivery in desmoplastic tumors. *Cancer Res*, *77*(3), 719-731. doi: 10.1158/0008-5472.CAN-16-0866

[23] Zhang S, Fang W, Zhou S, Zhu D, Chen R, Gao X, Li Y, Liu J, Wang H, Xu Y, Zhang Q, Sun J (2023). Single cell transcriptomic analyses implicate an immunosuppressive tumor microenvironment in pancreatic cancer liver metastasis. *Nat Commun*, *14*(1), 5123. doi: 10.1038/s41467-023-40727-7

[24] Thommen DS, Schumacher TN (2018). T cell dysfunction in cancer. *Cancer Cell*, *33*(4), 547-562. doi: 10.1016/j.ccell.2018.03.012

[25] Hegde S, Krisnawan VE, Herzog BH, Zuo C, Breden MA, Knolhoff BL, Hogg GD, Teixeiro E, Alspach E, Schmidt SV, Richer MJ, Hiam-Galvez KJ, Gillanders WE, DeNardo DG (2020). Dendritic cell paucity leads to

dysfunctional immune surveillance in pancreatic cancer. *Cancer Cell*, *37*(3), 289-307.e9. doi: 10.1016/j.ccell.2020.02.008

[26] Rudloff MW, Zumbo P, Favret NR, Roetman JJ, Detrés Román CR, Erwin MM, Zhai Y, Fernandez N, Su X, Paley MA, Betts MR, Wherry EJ (2023). Hallmarks of CD8+ T cell dysfunction are established within hours of tumor antigen encounter before cell division. *Nat Immunol*, *24*(9), 1527-1539. doi: 10.1038/s41590-023-01578-y

[27] Mahadevan KK, McAndrews KM, LeBleu VS, Yang S, Lyu H, Li B, Tiriac H, Bailey JM, Laheru DA, Hruban RH, Iacobuzio-Donahue CA, Wood LD, Kimmelman AC (2023). KRASG12D inhibition reprograms the microenvironment of early and advanced pancreatic cancer to promote FAS-mediated killing by CD8+ T cells. *Cancer Cell*, *41*(9), 1606-1620.e8. doi: 10.1016/j.ccell.2023.07.002

[28] Zhao Y, Tang J, Jiang K, Liu SY, Aicher A, Heeschen C (2023). Liquid biopsy in pancreatic cancer: current perspective and future outlook. *Biochim Biophys Acta Rev Cancer*, *1878*(3), 188868. doi: 10.1016/j.bbcan.2023.188868

[29] Lin YN, Schmidt MO, Sharif GM, Vietsch EE, Kiliti AJ, Barefoot ME, Shah YM, Lyssiotis CA, Frankel TL, Bednar F (2022). Impaired CXCL12 signaling contributes to resistance of pancreatic cancer subpopulations to T cell-mediated cytotoxicity. *Oncoimmunology*, *11*(1), 2027136. doi: 10.1080/2162402X.2022.2027136

[30] Schietinger A, Philip M, Krisnawan VE, Chiu EY, Delrow JJ, Basom RS, Lauer P, Brockstedt DG, Knoblaugh SE, Houghton AN, Greenberg PD (2016). Tumor-specific T cell dysfunction is a dynamic antigen-driven differentiation program initiated early during tumorigenesis. *Immunity*, *45*(2), 389-401. doi: 10.1016/j.immuni.2016.07.011

[31] Jacenik D, Lebish EJ, Beswick EJ (2023). MK2 drives progression of pancreas and colon cancers by suppressing CD8+ T cell cytotoxic function and is a potential immunotherapy target. *Front Immunol*, *14*, 1212100. doi: 10.3389/fimmu.2023.1212100

[32] Ajina R, Weiner LM (2020). T-cell immunity in pancreatic cancer. *Pancreas*, *49*(8), 1014-1023. doi: 10.1097/MPA.0000000000001621

[33] Steele NG, Carpenter ES, Kemp SB, Sirihorachai VR, Delrosario L, Lazarus J, Amir EAD, Gunchick V, Espinoza C, Bell JS, Harris KT, Lima F, Irizarry-Negron V, Pagliarini R, Wang J, Zaw Thin M, Yuan J, Bednar F, Pasca di Magliano M (2020). Multimodal mapping of the tumor and peripheral blood immune landscape in human pancreatic cancer. *Nat Cancer*, *1*(11), 1097-1112. doi: 10.1038/s43018-020-00121-4

[34] Sang W, Zhou Y, Chen H, Yu C, Dai L, Liu Z, Zhang X, Li Y, Wang Q, Sun R, Zhao J, Chen Y, Liu J, Xu K, Wang H, Zhang L (2024). Receptor-interacting protein kinase 2 is an immunotherapy target in pancreatic cancer. *Cancer Discov*, *14*(2), 326-347. doi: 10.1158/2159-8290.CD-23-0584

[35] O'Donnell JS, Teng MWL, Smyth MJ (2019). Cancer immunoediting and resistance to T cell-based immunotherapy. *Nat Rev Clin Oncol*, *16*(3), 151-167. doi: 10.1038/s41571-018-0142-8

[36] Han J, Khatwani N, Searles TG, Turk MJ, Angeles CV (2020). Memory CD8+ T cell responses to cancer. *Semin Immunol*, *49*, 101435. doi: 10.1016/j.smim.2020.101435

[37] Philip M, Schietinger A (2022). CD8+ T cell differentiation and dysfunction in cancer. *Nat Rev Immunol*, *22*(4), 209-223. doi: 10.1038/s41577-021-00574-3

[38] van der Leun AM, Thommen DS, Schumacher TN (2020). CD8+ T cell states in human cancer: insights from single-cell analysis. *Nat Rev Cancer*, *20*(4), 218-232. doi: 10.1038/s41568-019-0235-4

[39] Chen H, Zhuo Q, Ye Z, Xu X, Ji S (2021). Organoid model: a new hope for pancreatic cancer treatment? *Biochim Biophys Acta Rev Cancer*, *1875*(1), 188466. doi: 10.1016/j.bbcan.2020.188466

[40] Chen K, Wang Y, Hou Y, Wang Q, Long D, Liu X, Zhang H, Li J, Zhou Y, Xu L, Sun X, Zhao Q (2022). Single cell RNA-seq reveals the CCL5/SDC1 receptor-ligand interaction between T cells and tumor cells in pancreatic cancer. *Cancer Lett*, *545*, 215834. doi: 10.1016/j.canlet.2022.215834

[41] Qin C, Yang G, Yang J, Ren B, Wang H, Chen G, Zhao F, You L, Wang W, Zhao Y (2020). Metabolism of pancreatic cancer: paving the way to better anticancer strategies. *Mol Cancer*, *19*(1), 50. doi: 10.1186/s12943-

020-01169-7

[42] Zhang Z, Wang ZX, Chen YX, Wu HX, Yin L, Zhao Q, Li J, Zhang Y, Liu H, Sun S, Xu L (2022). Integrated analysis of single-cell and bulk RNA sequencing data reveals a pan-cancer stemness signature predicting immunotherapy response. *Genome Med*, *14*(1), 45. doi: 10.1186/s13073-022-01050-w

[43] Ashok G, Miryala SK, Saju MT, Anbarasu A, Ramaiah S (2022). FN1 encoding fibronectin as a pivotal signaling gene for therapeutic intervention against pancreatic cancer. *Mol Genet Genomics*, *297*(6), 1565-1580. doi: 10.1007/s00438-022-01943-w

[44] Rojas LA, Sethna Z, Soares KC, Olcese C, Pang N, Patterson E, Lihm J, Ceglia N, Guasp P, Chu A, Snyder A, Houghton A, Wolchok JD, Merghoub T, Balachandran VP (2023). Personalized RNA neoantigen vaccines stimulate T cells in pancreatic cancer. *Nature*, *618*(7963), 144-150. doi: 10.1038/s41586-023-06063-y

[45] Bhatt D, Kang B, Sawant D, Zheng L, Perez K, Huang Z, Vella JL, Veatch JR, Riddell SR, Zheng H, Robins HS (2021). STARTRAC analyses of scRNAseq data from tumor models reveal T cell dynamics and therapeutic targets. *J Exp Med*, *218*(6), e20201329. doi: 10.1084/jem.20201329

[46] Wu J (2022). Pancreatic cancer-derived exosomes promote the proliferation, invasion, and metastasis of pancreatic cancer by the miR-3960/TFAP2A axis. *J Oncol*, 2022, 3590326. doi: 10.1155/2022.3590326

[47] Hamidi-Sofiani V, Rakhshi R, Moradi N, Zeynali P, Nakhaie M, Behboudi E (2022). Oncolytic viruses and pancreatic cancer. *Cancer Treat Res Commun*, *31*, 100563. doi: 10.1016/j.ctarc.2022.100563