Available Online: 25 December 2025 DOI: 10.54254/3049-5458/2025.30934

Bioinformatics analysis for pancreatic cancer to identify risk genes based on blood samples

Tianyang Feng

University of California, San Diego, USA

t4feng@ucsd.edu

Abstract. Pancreatic cancer remains a highly lethal malignancy with limited early detection methods and therapeutic options. To better understand its molecular underpinnings, we performed a comprehensive single-cell RNA-sequencing analysis of Peripheral Blood Mononuclear Cells (PBMCs) from 16 pancreatic cancer patients and 4 healthy controls. After integrating datasets using Harmony to correct for batch effects, we annotated immune cells across three levels of resolution, identifying Level 2 as optimal for downstream analysis. Differential gene expression and pathway enrichment analyses revealed distinct transcriptional signatures across immune subsets, including $CD4^+$ T cells, monocytes, NK cells, and regulatory T cells. Notably, monocytes exhibited pronounced pro-inflammatory shifts, while unexpected expression of pancreatic enzyme-related genes was observed in erythroid cells and Tregs, suggesting potential contamination or stress-induced pathways. Our study provides a high-resolution map of the systemic immune landscape in pancreatic cancer and highlights candidate cellular and transcriptional mechanisms that may contribute to disease pathology, offering potential targets for future diagnostic and therapeutic strategies.

Keywords: pancreatic cancer, scRNA-seq, PBMCs, immune landscape, bioinformatics analysis

1. Introduction

Pancreatic cancer is one of the most aggressive and lethal malignancies, it can develop 57,600 new cases and cause 47,050 deaths in one year in US, which is 8% of the total cancer death [1]. Despite advances in cancer research, the early detection technic of pancreatic cancer are not advance as well and still need further develop [2]. Study its molecular mechanisms could lead to targeted therapies, including mutations in genes BRCA1, BRCA2, PALB2, and CDKN2A [3]. The pancreatic cancer is highly resistant to conventional treatments, that displayed by its dense stroma and immunosuppressive nature [4]. Research of the risk factors, such as physical activity, obesity, and height, may also help in prevention and early intervention strategies [5]. Furthermore, improving treatment options could control the physical burden (pain, fatigue, weight loss) and psychological distress on patients and their families [6]. Finally, advancing pancreatic cancer research has the potential to save lives and reduce healthcare costs [7].

Single cell sequencing could provide resolution of cellular heterogeneity by analysis individual cells, and overcome the limitations of bulk sequencing [8]. This technique could identify rare cell populations that are often hided in bulk data, which will provide deeper biological insights [9]. Single cell sequencing reveals

multiple gene expression patterns across different cell types, enhancing the understand of developmental processes [10]. It allows precise characterization of tumor microenvironments which could help cancer research and personalized medicine [11]. This technique also promote the discovery of new cell subtypes, refining cell-type classifications in complex tissues [12]. Through the clonal evolution in cancers, single cell sequencing provides critical insights into disease progression and treatment resistance [13]. Additionally, the single cell sequencing improves the study of immune cell diversity and function [14].

Single-cell sequencing provide new insights about oncology by enabling the dissection of intratumoral heterogeneity at high resolution and display distinct subpopulations of cancer cells within tumors [15]. Recent studies have applied single cell sequencing to study clonal evolution in real-time, discovering mechanisms of therapy resistance in cancers such as melanoma and lung adenocarcinoma [16]. In 2023, researchers used single-cell RNA sequencing to identify rare circulating tumor cells in liquid biopsies which contribute to early cancer detection [17]. Single cell sequencing has also contributed to the tumor microenvironment by displaying how immune cell interactions influence immunotherapy response in breast and colorectal cancers [18]. Advances in multi-omics single-cell approaches have mapped epigenetic and transcriptomic changes causing metastasis in pancreatic ductal adenocarcinoma [19]. A 2022 study used single-cell ATAC-seq to display chromatin accessibility patterns linked to drug resistance in glioblastoma [20]. Recently, spatial transcriptomics combined with single cell sequencing has enabled 3D mapping of tumor-immune cell interactions [21].

2. Results

2.1. Integration of single-cell datasets from multiple donors

To evaluate the consistency of single-cell transcriptomic profiles across individuals, we generated UMAP embeddings of single-cell RNA-sequencing data from 20 donors, including 4 healthy controls and 16 Peripheral Blood Mononuclear Cell (PBMC) samples. As shown in Figure 1a, prior to batch correction, cells clustered predominantly by donor identity, indicating the presence of substantial inter-sample variability likely due to technical or batch effects.

To address this issue, we applied Harmony method to align the datasets across samples. After integration (Figure 1b), cells no longer segregated by donor, suggesting that batch effects were effectively minimized. Importantly, the major cellular clusters remained intact, indicating that biological variation was preserved during the correction process. These results demonstrate that the integration method successfully harmonized transcriptomic profiles across individuals while maintaining meaningful biological structure in the data.

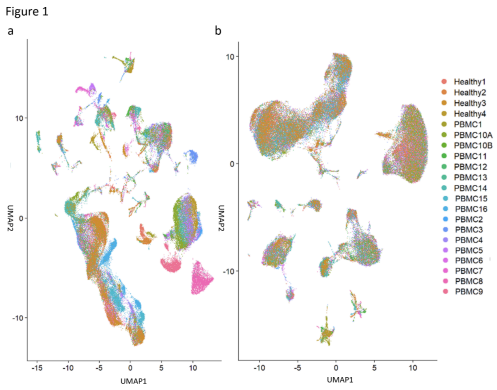

Figure 1. Integration of single-cell transcriptomic data from healthy and PBMC donors. (a) UMAP projection of single-cell RNA-sequencing data from 20 samples including 4 healthy controls (Healthy1–4) and 16 PBMC donor samples (PBMC1–16), colored by individual sample. Cells from each sample are distributed across clusters, indicating that cell types are preserved across individuals. (b) UMAP projection of the same dataset after integration using Harmony method, demonstrating improved alignment across donors. The absence of sample-specific clustering in panel (b) suggests successful correction for batch effects.

2.2. Immune cell composition and annotation resolution

To comprehensively define the immune landscape, we employed reference-based annotation of single-cell RNA sequencing data across three levels of granularity (Figure 2). Dimensionality reduction by UMAP revealed a well-structured transcriptional continuum of immune cell types.

At Level 1 (Figure 2a), major immune lineages were clearly resolved, including B cells, CD4 T cells, CD8 T cells, monocytes, NK cells, Dendritic Cells (DCs), and a group of undefined or rare cell types labeled as "other." Increasing the resolution to Level 2 (Figure 2b) allowed discrimination of key functional subsets. For instance, $CD14^+$ and $CD16^+$ monocytes emerged as distinct populations, as did naive vs memory B cells, and central memory (TCM) vs effector memory (TEM) T cells within both $CD4^+$ and $CD8^+$ compartments. Regulatory T cells (Tregs) and dendritic cell subtypes (cDC1, cDC2) also became distinguishable at this level.

At the highest Level 3 resolution (Figure 2c), we identified extensive heterogeneity within immune lineages. B cells were subclassified into naive, memory, and isotype-specific (kappa/lambda) groups. $CD4^+$ T cells were further divided into naive, central memory (TCM), effector memory (TEM), proliferating, and cytotoxic CD4 CTL subsets. Similarly, $CD8^+$ T cells showed diverse subtypes including naïve, proliferating, and six distinct TEM clusters. Rare but functionally significant populations such as MAIT cells, gdT cells, Innate Lymphoid Cells (ILCs), and NK CD56bright cells were also detected.

This hierarchical annotation strategy enabled high-resolution mapping of immune cell diversity and provides a framework for downstream analyses of functional states, disease-associated changes, and transcriptional dynamics across immune subsets. Additionally, as we consider the all three levels of resolution, level1 is over generalize for further analysis, while level3 is excessively detailed that some cell type only contain less than ten sample, level2 is the most suitable level for further analysis.

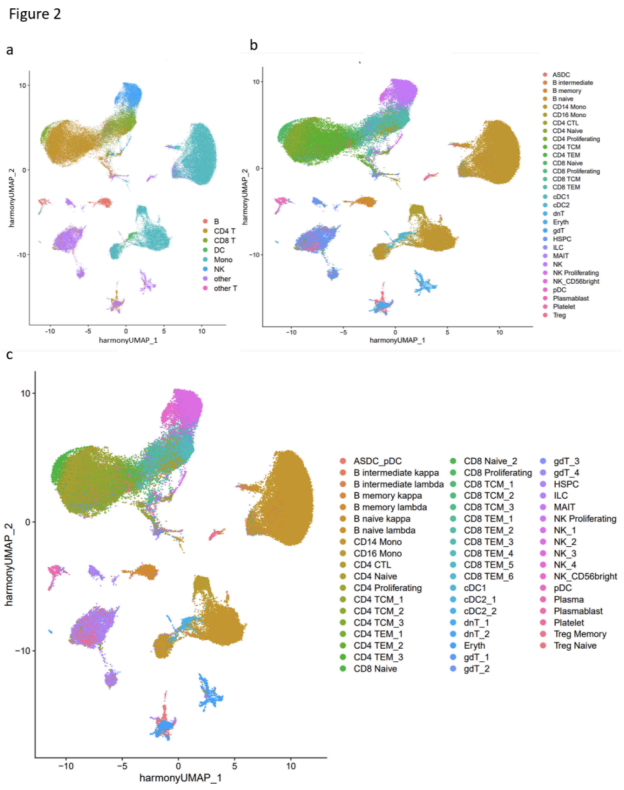

Figure 2. Immune cell type annotation across increasing resolution levels. (a–c) UMAP plots depicting immune cell populations identified using a reference-based cell type prediction method across three levels of annotation granularity: (a) Level 1: Broad immune cell types including B cells, CD4 T cells, CD8 T cells, dendritic cells (DC), monocytes (Mono), natural killer cells (NK), and others. (b) Level 2: Intermediate resolution showing subsets within major lineages such as $CD14^+$ vs $CD16^+$ monocytes, naive vs memory B and T cells, dendritic cell subsets (cDC1, cDC2), and regulatory T cells. (c) Level 3: High-resolution annotation identifying finer immune cell subtypes, including B cell isotypes (e.g., kappa vs lambda), multiple effector and memory subsets within CD4 and CD8 T cells (e.g., TCM, TEM, CTL), proliferative states, NK subtypes, gdT cells, plasmablasts, and rare populations such as MAIT and ILC. Each point represents a single cell colored by its predicted cell type. UMAP embedding was used for dimensionality reduction and visualization of transcriptional similarity.

2.3. Transcriptomic changes across immune cell subsets

To identify transcriptional changes across different immune cell populations, we performed differential gene expression analysis in eleven major immune subsets, and show five of them here: $CD4^+$ Naive T cells, $CD4^+$ Central Memory (TCM), $CD4^+$ Effector Memory (TEM), $CD14^+$ monocytes, and $CD16^+$ monocytes (Figure 3).

In CD4 Naive T cells (Figure 3a), several ribosomal protein genes (e.g., RPS17, RPL21, RPS12) and mitochondrial genes (MT-RNR2-L12, MT-RNR2-L8) were upregulated, along with the long non-coding RNA XIST, suggesting increased translational and mitochondrial activity. Downregulated genes included IF144L, LAIR2, and VAV3, possibly indicating reduced interferon signaling. CD4 TCM cells (Figure 3b) displayed a similar pattern, with upregulation of XIST, NBEAL1, and ZNF90, while genes such as UPK3BL and AC005082.1 were significantly downregulated. Notably, HLA-DQA2 was also among the top upregulated

genes, suggesting enhanced antigen presentation capacity. In contrast, CD4 TEM cells (Figure 3c) showed fewer differentially expressed genes. However, PRR7-AS1 and MTSS1 were significantly downregulated, indicating selective transcriptomic changes in effector memory subsets.

CD14^{+} monocytes (Figure 3d) exhibited robust transcriptional changes, with several innate immune genes (G0S2, FAM157B, DEFA5) significantly upregulated. Downregulated genes included GNLY, GZSH, and LAIR2, suggesting altered cytotoxic and immunoregulatory activity in this population. While CD16^{+} monocytes (Figure 3e) displayed extensive gene expression changes, with numerous interferon-stimulated genes (e.g., IFIT1, IFIT3, IFI44L, LAIR2) significantly downregulated. Meanwhile, genes associated with inflammation and stress response, including XIST, PER1, and HLA-DQA2, were upregulated, suggesting a shift toward a pro-inflammatory phenotype.

Together, these results demonstrate that immune cell subsets exhibit distinct transcriptional responses, with monocytes (particularly CD16^{+} monocytes) showing the most pronounced gene expression shifts, likely reflecting differential roles in immune activation or disease pathology.

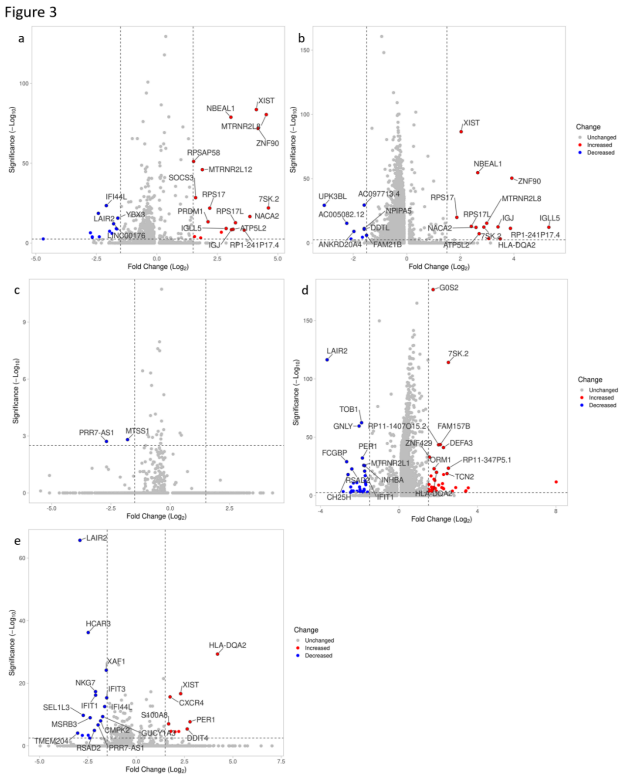

Figure 3. Differential gene expression in immune cell subsets. (a–e) Volcano plots showing differential gene expression profiles in five immune cell subsets: (a) CD4 Naive T cells, (b) CD4 Central Memory (TCM) T cells, (c) CD4 Effector Memory (TEM) T cells, (d) CD14^{+} monocytes, and (e) CD16^{+} monocytes. Each point represents a gene, with $\log_2$ fold change (x-axis) plotted against $-\log_{10}$ adjusted p-value (y-axis). Genes with statistically significant increases in expression (adjusted $p < 0.05$, fold change > 2) are highlighted in red, while significantly decreased genes are shown in blue. Non-significant genes are shown in grey. Selected genes of interest are labeled. Vertical dashed lines indicate ± 1 $\log_2$ fold change, and horizontal dashed lines indicate significance thresholds.

2.4. Transcriptional changes in erythroid, progenitor, and rare immune cell populations

We next examined differential gene expression in additional immune and hematopoietic populations, including erythroid cells, Hematopoietic Stem and Progenitor Cells (HSPCs), NK cells, plasmablasts, platelets, and regulatory T cells (Tregs) (Figure 4).

In erythroid cells (Figure 4a), there was a striking downregulation of multiple pancreatic enzyme-related genes such as CELA3A, PRSS3, CPA1, REG1A, CTRB1, and AMY2A, indicating possible contamination from digestive gene signatures or deconvolution artifacts. Only a few genes showed increased expression, with modest fold changes. HSPCs have a different pattern. they (Figure 4b) showed a clear transcriptional skew, with decreased expression of genes like SNCA, AQP5, STXBP1, and HEPACAM, while ribosomal gene RPL22L1 and TM4SF4 were significantly upregulated. These changes may reflect shifts in cell cycle, differentiation, or stress responses in progenitor states.

In NK cells (Figure 4c), decreased expression was observed for immune signaling and regulatory genes, including RAB34, HRASLS5, SOCS1, and XIRP1. Increased expression of genes such as XIST, TMEM2, and HLA-DQA2 suggests altered activation or regulatory state in NK subsets. Plasmablasts (Figure 4d) exhibited downregulation of multiple ion channel and transport genes including AQP2, FXYD2, KCNJ16, and ATP6V1C1, possibly indicating changes in membrane potential or metabolism. Upregulated genes were fewer but included immune effectors such as MAL, TMEM213, and TMEM52B.

Platelets (Figure 4e) demonstrated the most extreme transcriptional changes, with a broad and significant downregulation of genes involved in signal transduction (SNCA, CXCL17, MSLN, SLC6A14) and platelet function. In contrast, upregulated genes included RPL18A, RPS3A, and KRT18, suggesting shifts in translation or structural components. Finally, in Tregs (Figure 4f), we observed robust upregulation of pancreatic enzyme-related genes such as SPINK1, GP2, REG1A, PRSS3, CELA2A, and CTRB1, mirroring the pattern observed in erythroid cells. This unusual expression profile may indicate contamination, misclassification, or shared stress-induced transcriptional programs. Concurrently, classical Treg markers such as FOXP3 and IL2RA were not prominently altered, raising questions about subset identity or annotation accuracy.

Overall, these findings reveal pronounced transcriptional shifts in several rare or specialized immune populations. Notably, platelet and erythroid cells showed the greatest downregulation, while Tregs and HSPCs demonstrated unexpected upregulation of tissue-specific or stress-responsive genes, warranting further investigation.

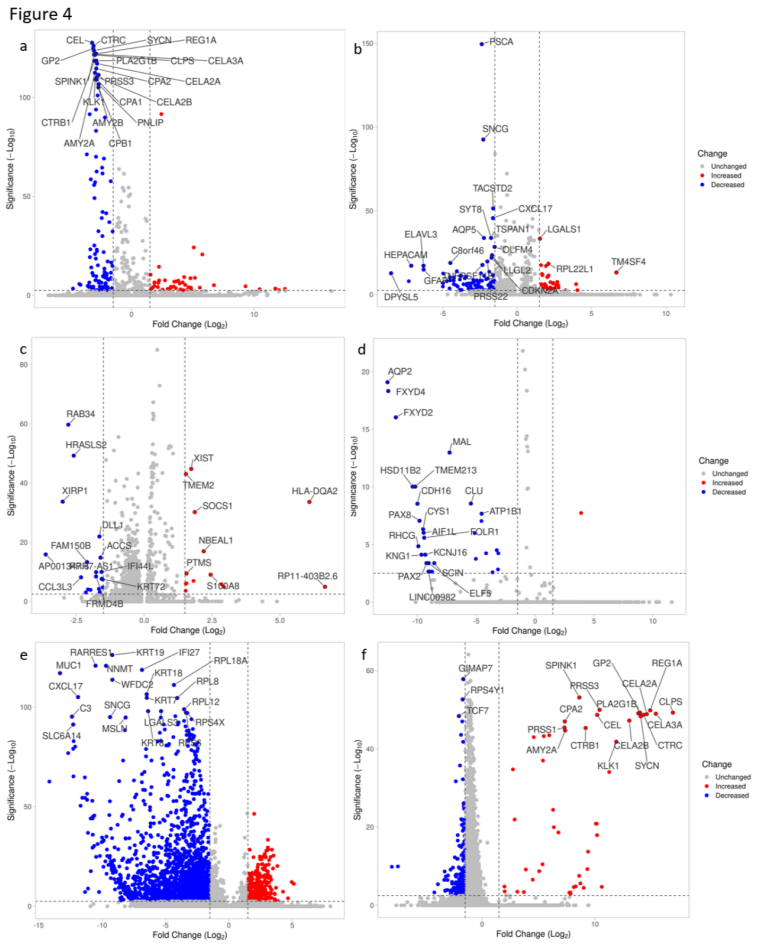

Figure 4. Differential gene expression in non-T cell immune subsets. (a–f) Volcano plots showing differential gene expression in: (a) Erythroid cells (Eryth), (b) Hematopoietic stem and progenitor cells (HSPC), (c) Natural killer cells (NK), (d) Plasmablasts, (e) Platelets, and (f) Regulatory T cells (Treg).Each point represents a gene, plotted by $\log_2$ fold change (x-axis) versus $-\log_{10}$ adjusted *p*-value (y-axis). Genes with significantly increased expression are shown in red, decreased in blue, and unchanged in gray. Labeled genes indicate those with the most significant changes. Vertical dashed lines represent ± 1 $\log_2$ fold change, and the horizontal dashed line marks the significance threshold (adjusted $p < 0.05$).

2.5. Functional enrichment of differential gene expression highlights distinct pathway signatures in immune cell subsets

To gain insights into the biological processes underlying the observed transcriptional changes, we performed pathway enrichment analysis on the differentially expressed genes across several major immune subsets (Figure 5).

In CD4 Naive T cells (Figure 5a), enriched pathways included ribosome-related functions, PD-L1/PD-1 checkpoint signaling, T cell receptor signaling, and spliceosome activity. Infection-related pathways such as coronavirus disease (COVID-19) and viral infections (e.g., herpes simplex, influenza) were also enriched, alongside stress-response and apoptosis-related signaling. These signatures suggest an early activation state with enhanced translational and immune checkpoint activity. Similarly, CD4 TCM cells (Figure 5b) showed a strong enrichment in T cell receptor signaling, Th1 and Th2 differentiation, spliceosome, and protein

processing in the endoplasmic reticulum. Several infection-related pathways (e.g., hepatitis B, Salmonella, HIV, and COVID-19) were also present. The enrichment of lipid metabolism and neurodegenerative disease pathways may reflect broader metabolic remodeling in central memory subsets. In contrast, CD4 TEM cells (Figure 5c) demonstrated enrichment in diverse signaling pathways, including lipolysis regulation, estrogen signaling, MAPK signaling, and cytokine-cytokine receptor interaction, as well as associations with disease-specific processes such as colorectal cancer and Chagas disease. These findings highlight a more metabolically active and inflammation-prone phenotype compared to naive or central memory T cells.

$CD14^+$ monocytes (Figure 5d) exhibited robust enrichment in innate immune signaling pathways such as the NOD-like receptor signaling pathway, chemokine signaling, FoxO signaling, and responses to bacterial and viral infections (e.g., Yersinia, Escherichia coli, COVID-19). Notably, osteoclast differentiation and endocytosis pathways were also prominent, consistent with the phagocytic and antigen-presenting roles of monocytes. $CD16^+$ monocytes (Figure 5e) showed the broadest and most diverse pathway enrichment, including strong associations with Chagas disease, leishmaniasis, relaxin and estrogen signaling, and platelet activation. Noteworthy immune pathways included Toll-like receptor signaling, cytokine signaling, and neutrophil extracellular trap formation, as well as several metabolic and cancer-related processes. This reflects the highly inflammatory and tissue-migratory nature of $CD16^+$ monocytes.

Collectively, these analyses reveal cell type-specific pathway activation patterns, with memory and effector T cells enriched for cytokine and metabolic processes, while monocytes engage diverse inflammatory, infectious, and innate immune pathways. These distinctions support functional specialization within the immune landscape and provide mechanistic insights into their differential responses.

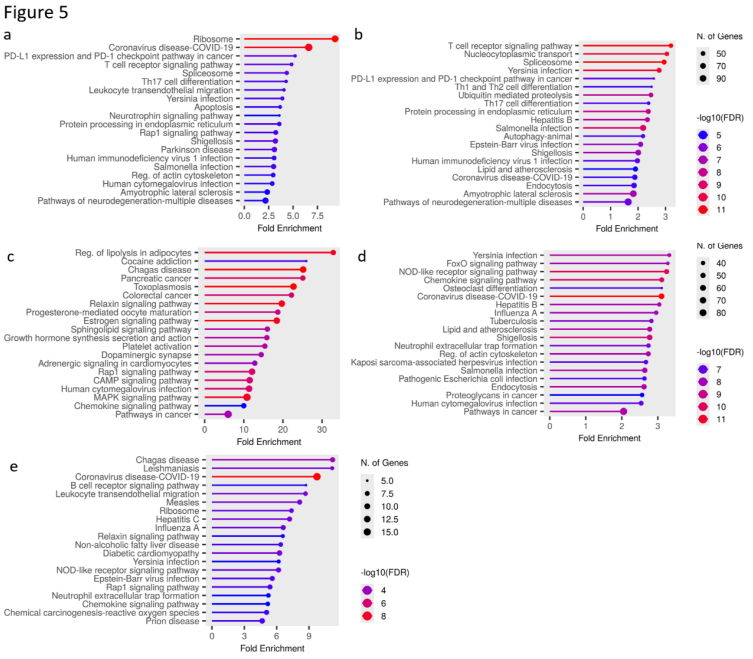

Figure 5. Pathway enrichment analysis in differentially expressed genes across immune cell subsets. (a–e) Bar plots showing pathway enrichment analysis of differentially expressed genes in: (a) CD4 Naive T cells, (b) CD4 Central Memory T cells (TCM), (c) CD4 Effector Memory T cells (TEM), (d) $CD14^+$ Monocytes, and (e) $CD16^+$ Monocytes. The x-axis indicates fold enrichment for each pathway, while the color scale reflects the -$\log_{10}$(FDR) value (adjusted false discovery rate). The size of the dots corresponds to the number of genes associated with each pathway. Top significantly enriched pathways include immune signaling, infection-related processes, cancer-related pathways, and metabolic functions, with variation depending on cell type.

2.6. Functional enrichment reveals shared and distinct pathway signatures in non-T immune subsets

To further investigate the biological relevance of gene expression changes in non-T immune subsets, we performed pathway enrichment analyses across six distinct populations (Figure 6). This revealed both shared and cell type–specific functional signatures.

In erythroid cells (Figure 6a), top-enriched pathways included ribosome, COVID-19, and various metabolic and protein processing pathways (e.g., glycine/serine/threonine metabolism, fat digestion and absorption, oxidative phosphorylation). These signatures reflect the high biosynthetic and metabolic demands of erythroid precursors and potential stress responses. HSPCs (Figure 6b) also showed enrichment for ribosome, oxidative phosphorylation, and estrogen signaling, suggesting active transcriptional and metabolic remodeling. Notably, enrichment in dopaminergic and cell adhesion pathways may reflect migration and niche retention mechanisms during hematopoiesis.

NK cells (Figure 6c) displayed a more immunologically focused enrichment profile, including B cell receptor signaling, T cell receptor signaling, Fc gamma R-mediated phagocytosis, and chemokine signaling pathways. Viral and bacterial infection pathways (Yersinia, COVID-19, Escherichia coli) were also prominent, reflecting the cytotoxic and pathogen-responsive role of NK cells. In plasmablasts (Figure 6d), enrichment centered on oxidative phosphorylation, complement and coagulation cascades, thyroid hormone synthesis, and endocrine-related signaling. Several metabolic and cardiac-associated pathways were also represented, potentially reflecting the high secretory and energy demands of antibody-producing cells.

Platelets (Figure 6e) exhibited strong enrichment in oxidative phosphorylation, protein processing, and neurodegenerative disease pathways. Notably, ribosome and Parkinson's disease pathways were highly enriched, consistent with platelet roles in inflammation, mitochondrial metabolism, and systemic responses to disease. Finally, Tregs (Figure 6f) showed high enrichment for ribosome, proteasome, oxidative phosphorylation, and COVID-19-related pathways. Additionally, several immune regulation and cancer-related pathways, including salmonella infection, chemical carcinogenesis, and T cell leukemia virus infection, were among the top pathways. These results suggest that Tregs may undergo transcriptional shifts indicative of enhanced protein turnover, immune modulation, or stress signaling.

Overall, this analysis reveals that ribosome and oxidative phosphorylation pathways are commonly enriched across multiple cell types, reflecting general metabolic activation, while immune-specific and cell-type-defining pathways emerge in subsets like NK cells and plasmablasts. These results underscore the functional heterogeneity and specialized roles of non-T immune cells under the studied condition.

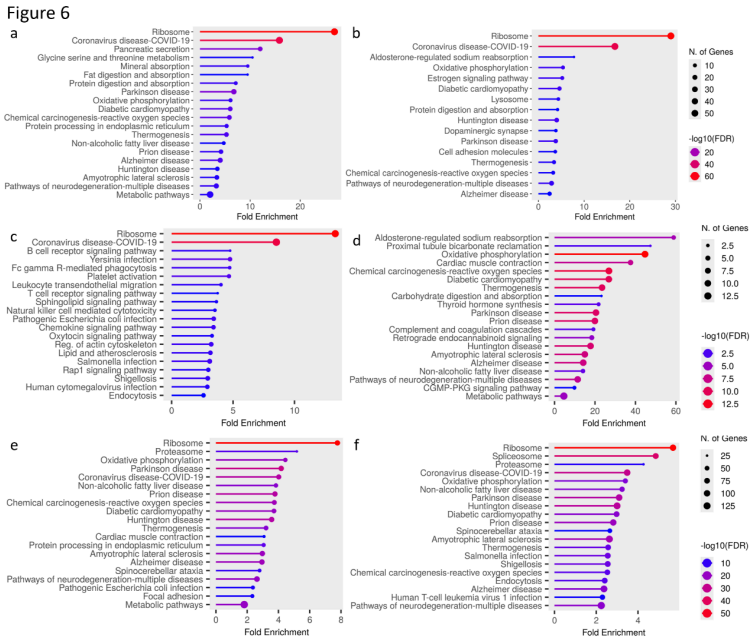

Figure 6. Pathway enrichment analysis in non-T immune and hematopoietic cell subsets. (a–f) Bar plots showing pathway enrichment analysis of differentially expressed genes in: (a) Erythroid cells (Eryth), (b) Hematopoietic Stem and Progenitor Cells (HSPC), (c) Natural Killer (NK) cells, (d) Plasmablasts, (e) Platelets, and (f) Regulatory T cells (Treg). The x-axis shows fold enrichment of significantly enriched pathways. Dot color indicates the $-\log_{10}$(FDR) significance level, while dot size corresponds to the number of genes contributing to each pathway. Ribosome and COVID-19-related pathways were consistently among the top-enriched across multiple cell types, while cell-type-specific enrichment patterns were also observed.

3. Discussion

In this study, we employed single-cell RNA sequencing of PBMCs to describe the systemic immune landscape associated with pancreatic cancer [22]. Our integrated analysis of samples from 16 patients and 4 healthy controls revealed profound and cell-type-specific transcriptional alterations, providing insights into the potential mechanisms of immune dysregulation in this devastating disease [23]. An important finding was the successful application of the Harmony algorithm to integrate multi-donor scRNA-seq data, effectively control technical batch effects while preserving biological heterogeneity [24]. Our hierarchical cell annotation strategy further mentions the continuum of immune cell states, with Level 2 resolution providing a perfect balance between biological insight and analytical power [25].

The differential gene expression analysis displayed significant transcriptional shifts across nearly all major immune lineages [26]. Most strikingly, $CD16^+$ monocytes exhibited a broad downregulation of interferon-stimulated genes alongside an upregulation of pro-inflammatory markers [27]. This signature is consistent with a previously described immunosuppressive yet inflammation-prone monocyte state that can contribute to tumor progression [28]. The pronounced pathway enrichment for Toll-like receptor signaling in these cells further suggests their involvement in chronic inflammatory processes, a characteristic of the pancreatic cancer microenvironment [29].

Another interesting observation was the significant upregulation of pancreatic enzyme genes in Tregs and their downregulation in erythroid cells [30]. While this could come from technical artifacts such as ambient

RNA contamination [31], it may also reflect a novel, stress-induced transcriptional program. The fact that typical Treg markers were not altered raises important questions about the identity of this subset, ensuring validation with protein-level assays [32].

The functional enrichment analyses consistently highlighted metabolic reprogramming across multiple cell types [33]. The widespread enrichment of ribosome and oxidative phosphorylation pathways points to a system-wide bioenergetic demand in pancreatic cancer patients [34]. This aligns with the concept of cancer as a systemic disease that reprograms distal immune cell metabolism [35]. The strong enrichment of COVID-19 related pathways across subsets is also notable and may reflect a cancer-induced mimicry of antiviral responses [36].

Our study has several limitations that should be considered. First, the analysis was performed on blood samples, which may only partially reflect the dynamics of the tumor microenvironment [37]. Second, the functional implications of the observed transcriptional changes are inferred computationally and require mechanistic validation [38]. Finally, the lack of detailed clinical metadata limits our ability to correlate these immune signatures with patient outcomes [39].

In conclusion, our high-resolution single-cell atlas of the PBMC compartment in pancreatic cancer patients reveals a complex and coordinated rewiring of the systemic immune response [40]. We identify specific monocyte and Treg subsets with distinctive transcriptional profiles that may serve as novel cellular biomarkers or therapeutic targets for this lethal malignancy [41].

4. Methods

4.1. Data acquisition

The data used in this study are available from NCBI (GSE 155698).

4.2. Data loading

Load all packages including Seurat and Azimuth. Load all data and name them by Read10X and then CreateSeuratObject for each sample. Add condition to each object, which are cancer and healthy. Totally there 17 cancer data sets and 4 healthy data sets.

4.3. Quality control

For each object go through:

```
obj[["percent.mt"]] <- PercentageFeatureSet(obj, pattern = "^MT-")
obj[['RNA']]
obj <- NormalizeData(obj)
obj<- FindVariableFeatures(obj)
obj<- ScaleData(obj)
```

4.4. Dimensional reduction

For each object:

```
obj <-RunPCA(obj)
obj <-FindNeighbors(obj,
```

```
dims=1:30,
reduction="pca")
obj <-FindClusters(obj,
resolution = 2)
obj <-RunUMAP(obj,
dims=1:30,
reduction="pca")
DimPlot(obj)
```

4.5. Cell type annotation using azimuth

Library Seuratdisk and data save as rds file. Azimuth annotation is on the website. After azimuth annotation, read those data and merged all the data sets as one:

```
PBMC.merged <-merge(PBMC1.obj, y=c(PBMC2.obj, PBMC3.obj, PBMC4.obj,
PBMC5.obj, PBMC6.obj, PBMC7.obj, PBMC8.obj, PBMC9.obj, PBMC10A.obj,
PBMC10B.obj, PBMC11.obj, PBMC12.obj, PBMC13.obj, PBMC14.obj, PBMC15.obj,
PBMC16.obj, Healthy1.obj, Healthy2.obj, Healthy3.obj, Healthy4.obj),
add.cell.ids = c("PBMC1", "PBMC2", "PBMC3", "PBMC4", "PBMC5", "PBMC6",
"PBMC7", "PBMC8", "PBMC9", "PBMC10A", "PBMC10B", "PBMC11", "PBMC12",
"PBMC13", "PBMC14", "PBMC15", "PBMC16","Healthy1", "Healthy2", "Healthy3",
"Healthy4"),
project = "PBMCintergrated")
PBMC.merged <- SCTransform(PBMC.merged)
PBMC.merged <- RunPCA(PBMC.merged)
```

4.6. Data integration: harmony

For the merged data, go through:

```
PBMC.merged <-NormalizeData(PBMC.merged)
PBMC.merged <-FindVariableFeatures(PBMC.merged)
PBMC.merged <-ScaleData(PBMC.merged)
PBMC.merged <-RunPCA(PBMC.merged)
PBMC.merged <-FindNeighbors(PBMC.merged,
dims=1:30,
reduction="pca")
PBMC.merged <-FindClusters(PBMC.merged,
resolution = 2,
cluster.name = "unintegratedcluster")
```

```
PBMC.merged <-RunUMAP(PBMC.merged,
dims=1:30,
reduction="pca",
reduction.name = "unintegratedUMAP")
DimPlot(PBMC.merged,
reduction = 'unintegratedUMAP',
group.by = "orig.ident")
```

4.7. Using harmony method to integrated data

For integrated data:

```
PBMC.Integrated <-IntegrateLayers(object = PBMC.merged,
method = HarmonyIntegration,
orig.reduction = "pca",
new.reduction = "harmony",
verbose = TRUE)
PBMC.Integrated <-FindNeighbors(PBMC.Integrated,
reduction="harmony",
dims=1:30)
PBMC.Integrated <-FindClusters(PBMC.Integrated,
resolution = 2,
cluster.name = "harmonycluster")
PBMC.Integrated <-RunUMAP(PBMC.Integrated,
reduction = "harmony",
dims=1:30,
reduction.name = "harmonyUMAP")
DimPlot(PBMC.Integrated,
reduction = 'harmonyUMAP',
group.by = "orig.ident")
```

4.8. Differential analysis

For the integrated data set:

```
PBMC.Integrated$celltype.condition <-
paste(PBMC.Integrated$predicted.celltype.l1,
PBMC.Integrated$condition,sep="_")
Idents(PBMC.Integrated)<-"celltype.condition"
```

```
table(PBMC.Integrated$celltype.condition)

DimPlot(PBMC.Integrated,reduction =
"harmonyUMAP",group.by="predicted.celltype.l1")

DimPlot shows in Figure 2.

For each cell type which have more than 100 samples:

Cell.integrated <- subset(PBMC.Integrated, subset = celltype.condition %in%
c("Cell_cancer","Cell_healthy"))

DefaultAssay(Cell.integrated) <- "RNA"

Cell.integrated <- JoinLayers(Cell.integrated)

Cell.de <- FindMarkers(Cell.integrated, ident.1="Cell_cancer",
ident.2="Cell_healthy")
```

4.9. Volcano plot

Volcano Plot is formed by using the VolcaNoseR website, and Fold Change is the x-axis while Significance is the y-axis. For the detailed settings, size of datapoints is 3, visibility of data is 1, fold change threshold is -1.5 to 1.5, significance threshold is 2.5, use threshold to annotate is changed (and significant), criterion for ranking hits are Manhattan distance, number of top hits are 20, color (unchanged, increased, decreased) is grey, red, blue, plot height is 600, and plot width is 800.

4.10. Go analysis

Go analysis is formed by the website ShinyGo 0.85. For the detailed settings, pathway database is KEGG, FDR cutoff is 0.05, #pathway to show is 20, pathway size from 2 to 5,000, sort pathway by fold enrichment, x-axis is fold enrichment, color is -log10(FDR), size is genes, front size is 12, circle size is 4, color: high is red and low is blue, chart type is lollipop, aspect ratio is 2, and plot theme is default.

References

[1] Rebecca L. Siegel, Siegel, R. L., Kimberly D. Miller, Miller, K. D., Ahmedin Jemal, & Jemal, A. (2020). Cancer statistics, 2020. *CA: A Cancer Journal for Clinicians, 70*(1), 7–30. https: //doi.org/10.3322/caac.21590

[2] Alison P. Klein, & Klein, A. P. (2021). Pancreatic cancer epidemiology: Understanding the role of lifestyle and inherited risk factors. *Nature Reviews Gastroenterology & Hepatology, 18*(7), 493–502. https: //doi.org/10.1038/s41575-021-00457-x

[3] David B. Zhen, Zhen, D. B., Kari G. Rabe, Rabe, K. G., Steven Gallinger, Gallinger, S., Sapna Syngal, Syngal, S., Ann G. Schwartz, Schwartz, A. G., Michael Goggins, Michael G. Goggins, Goggins, M., Ralph H. Hruban, Hruban, R. H., Michele L. Coté, Cote, M. L., Robert R. McWilliams, McWilliams, R. R., … Petersen, G. M. (2015). BRCA1, BRCA2, PALB2, and CDKN2A mutations in familial pancreatic cancer: A PACGENE study. *Genetics in Medicine, 17*(7), 569–577. https: //doi.org/10.1038/gim.2014.153

[4] Albrecht Neeße, Neesse, A., Christian Bauer, Bauer, C., Bauer, C., Daniel Öhlund, Öhlund, D., Matthias Lauth, Lauth, M., Malte Buchholz, Buchholz, M., Patrick Michl, Michl, P., David A. Tuveson, Tuveson, D. A., Thomas M. Gress, & Gress, T. M. (2019). Stromal biology and therapy in pancreatic cancer: Ready for clinical translation? *Gut, 68*(1), 159–171. https: //doi.org/10.1136/gutjnl-2018-316451

[5] Dominique S. Michaud, Michaud, D. S., Giovannucci, E., Willett, W. C., Colditz, G. A., Stampfer, M. J., & Fuchs, C. S. (2001). Physical Activity, Obesity, Height, and the Risk of Pancreatic Cancer. *JAMA, 286*(8), 921–

929. https: //doi.org/10.1001/jama.286.8.921

[6] Natalia Khalaf, Khalaf, N., Hashem B. El–Serag, El-Serag, H. B., Hannah R Abrams, Abrams, H. R., Aaron P. Thrift, & Thrift, A. P. (2020). Burden of Pancreatic Cancer-From Epidemiology to Practice. *Clinical Gastroenterology and Hepatology, 19*(5), 876–884. https: //doi.org/10.1016/j.cgh.2020.02.054

[7] Lola Rahib, Rahib, L., Rahib, L., Benjamin D. Smith, Benjamin D. Smith, Smith, B., Rhonda Aizenberg, Aizenberg, R., Allison Rosenzweig, Rosenzweig, A. B., Rosenzweig, A. B., Julie Fleshman, Fleshman, J., Fleshman, J. M., Lynn M. Matrisian, Matrisian, L. M., & Lynn M. Matrisian. (2014). Projecting Cancer Incidence and Deaths to 2030: The Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States. *Cancer Research, 74*(11), 2913–2921. https: //doi.org/10.1158/0008-5472.can-14-0155

[8] Fuchou Tang, Tang, F., Cátálin Bárbácioru, Barbacioru, C., Yangzhou Wang, Wang, Y., Ellen Nordman, Nordman, E., Clarence Lee, Lee, C., Nanlan Xu, Xu, N., Xiaohui Wang, Wang, X., John Bodeau, John Bodeau, Bodeau, J., Bodeau, J., Brian B. Tuch, … Surani, M. A. (2009). mRNA-Seq whole-transcriptome analysis of a single cell. *Nature Methods, 6*(5), 377–382. https: //doi.org/10.1038/nmeth.1315

[9] Evan Z. Macosko, Macosko, E. Z., Anindita Basu, Basu, A., Rahul Satija, Satija, R., James Nemesh, Nemesh, J., Karthik Shekhar, Shekhar, K., Melissa Goldman, Goldman, M., Itay Tirosh, Tirosh, I., Allison R. Bialas, Bialas, A. R., Nolan Kamitaki, Kamitaki, N., Emily M. Martersteck, … McCarroll, S. A. (2015). Highly Parallel Genome-wide Expression Profiling of Individual Cells Using Nanoliter Droplets. *Cell*. https: //doi.org/10.1016/j.cell.2015.05.002

[10] Cole Trapnell, & Trapnell, C. (2015). Defining cell types and states with single-cell genomics. *Genome Research, 25*(10), 1491–1498. https: //doi.org/10.1101/gr.190595.115

[11] Itay Tirosh, Tirosh, I., Benjamin Izar, Izar, B., Daniel J. Treacy, Treacy, D. J., John J. Trombetta, Trombetta, J. J., Asaf Rotem, Rotem, A., Asaf Rotem, Christopher Rodman, Rodman, C., Rodman, C., Mostaghimi, A., Christine G. Lian, Lian, C. G., Gabe J. Murphy, Murphy, G. F., … Garraway, L. A. (2016). Dissecting the multicellular ecosystem of metastatic melanoma by single-cell RNA-seq. *Science*. https: //doi.org/10.1126/science.aad0501

[12] Anoop P. Patel, Patel, A. P., Itay Tirosh, Tirosh, I., John J. Trombetta, Trombetta, J. J., Alex K. Shalek, Shalek, A. K., Gillespie, S. H., Shawn Gillespie, Gillespie, S. M., Hiroaki Wakimoto, Wakimoto, H., Daniel P. Cahill, Cahill, D. P., Brian V. Nahed, Nahed, B. V., Curry, W. J., William T. Curry, … Bernstein, B. E. (2014). Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma. *Science*. https: //doi.org/10.1126/science.1254257

[13] Charles Gawad, Gawad, C., Winston Koh, Koh, W., Stephen R. Quake, & Quake, S. R. (2016). Single-cell genome sequencing: Current state of the science. *Nature Reviews Genetics, 17*(3), 175–188. https: //doi.org/10.1038/nrg.2015.16

[14] Valentine Svensson, Svensson, V., Roser Vento-Tormo, Vento-Tormo, R., Sarah A. Teichmann, & Teichmann, S. A. (2018). Exponential scaling of single-cell RNA-seq in the past decade. *Nature Protocols, 13*(4), 599–604. https: //doi.org/10.1038/nprot.2017.149

[15] Sunny Z. Wu, Wu, S. Z., Ghamdan Al-Eryani, Al-Eryani, G., Daniel Roden, Roden, D. L., Simon Junankar, Junankar, S., Kate Harvey, Harvey, K., Alma Andersson, Andersson, A., Aatish Thennavan, Thennavan, A., Chenfei Wang, Wang, C., James R. Torpy, Torpy, J. R., Nenad Bartoniček, … Swarbrick, A. (2021). A single-cell and spatially resolved atlas of human breast cancers. *Nature Genetics, 53*(9), 1334–1347. https: //doi.org/10.1038/s41588-021-00911-1

[16] Charissa Kim, Kim, C., Ruli Gao, Gao, R., Emi Sei, Sei, E., Rachel Brandt, Brandt, R., Rachel Brandt, Johan Hartman, Hartman, J., Thomas Hatschek, Hatschek, T., Nicola Crosetto, Crosetto, N., Crosetto, N., Theodoros Foukakis, Foukakis, T., Nicholas E. Navin, & Navin, N. (2018). Chemoresistance Evolution in Triple-Negative Breast Cancer Delineated by Single-Cell Sequencing. *Cell, 173*(4), 879–893. https: //doi.org/10.1016/j.cell.2018.03.041

[17] Jeanne Tie, Tie, J., Joshua D. Cohen, Joshua D. Cohen, Kamel Lahouel, Lahouel, K., Serigne Lo, Serigne N. Lo, Yuxuan Wang, Yuxuan Wang, Suzanne Kosmider, Suzanne Kosmider, Rachel Wong, Wong, R., Jeremy Shapiro, Jeremy Shapiro, Margaret Lee, Margaret Lee, Sam Harris, … Peter Gibbs. (2022). Circulating Tumor DNA Analysis Guiding Adjuvant Therapy in Stage II Colon Cancer. *The New England Journal of Medicine*. https: //doi.org/10.1056/nejmoa2200075

[18] Lei Zhang, Zhang, L., Lei Zhang, Lei Zhang, Xin Yu, Xin Yu, Yu, X., Liangtao Zheng, Zheng, L., Yuanyuan Zhang, Zhang, Y., Yansen Li, Li, Y., Fang Qiao, Fang, Q., Ranran Gao, Gao, R., Boxi Kang, Kang, B., … Zhang, Z. (2018). Lineage tracking reveals dynamic relationships of T cells in colorectal cancer. *Nature, 564*(7735), 268–272. https: //doi.org/10.1038/s41586-018-0694-x

[19] Junya Peng, Junya, P., Bao-Fa Sun, Sun, B.-F., Chuanyuan Chen, Chen, C., Jiayi Zhou, Zhou, J.-Y., Yu-Sheng Chen, Chen, Y.-S., Hao Chen, Hao, C., Lulu Liu, Lulu, L., Dan Huang, Dan, H., Jialin Jiang, Jiang, J., Guanshen Cui, … Wu, W. (2019). Single-cell RNA-seq highlights intra-tumoral heterogeneity and malignant progression in pancreatic ductal adenocarcinoma. *Cell Research, 29*(9), 725–738. https: //doi.org/10.1038/s41422-019-0195-y

[20] Jin Ku Lee, Lee, J. K., Jiguang Wang, Wang, J., K. Jason, Sa, J. K., Erik Ladewig, Ladewig, E., Lee, H.-O., Hae Ock Lee, Lee, H. O., In Hee Lee, Lee, I.-H., Hyun Ju Kang, Hyun Ju Kang, Kang, H. J., Daniel Rosenbloom, Rosenbloom, D. I. S., Pablo G. Cámara, … Nam, D.-H. (2017). Spatiotemporal Genomic Architecture Informs Precision Oncology in Glioblastoma. *Nature Genetics, 49*(4), 594–599. https: //doi.org/10.1038/ng.3806

[21] Lu, L., Lambda Moses, Moses, L., Lior Pachter, & Pachter, L. (2020). Museum of Spatial Transcriptomics. *Nature Methods, 9*. https: //doi.org/10.7490/f1000research.1118091.1

[22] Rahul Satija, Satija, R., Jeffrey A. Farrell, Farrell, J. A., David Gennert, Gennert, D., Alexander F. Schier, Schier, A. F., Aviv Regev, & Regev, A. (2015). Spatial reconstruction of single-cell gene expression data. *Nature Biotechnology, 33*(5), 495–502. https: //doi.org/10.1038/nbt.3192

[23] Ashraful Haque, Haque, A., Jessica A. Engel, Engel, J. A., Sarah A. Teichmann, Teichmann, S. A., Tapio Lönnberg, & Lönnberg, T. (2017). A practical guide to single-cell RNA-sequencing for biomedical research and clinical applications. *Genome Medicine, 9*(1), 75–75. https: //doi.org/10.1186/s13073-017-0467-4

[24] Ilya Korsunsky, Korsunsky, I., Nghia Millard, Millard, N., Jean Fan, Fan, J., Kamil Slowikowski, Slowikowski, K., Fan Zhang, Zhang, F., Fan Zhang, Kevin Wei, Wei, K., Yuriy Baglaenko, Baglaenko, Y., Michael B. Brenner, Brenner, M. B., Po−Ru Loh, Loh, P.-R., … Raychaudhuri, S. (2019). Fast, sensitive and accurate integration of single-cell data with Harmony. *Nature Methods, 16*(12), 1289–1296. https: //doi.org/10.1038/s41592-019-0619-0

[25] Dvir Aran, Aran, D., Agnieszka Looney, Looney, A. P., Leqian Liu, Liu, L., Tian Xie, Wu, E., Valerie Fong, Fong, V., Austin Hsu, Hsu, A., Suzanna Chak, Chak, S., Ram P. Naikawadi, Naikawadi, R. P., Paul J. Wolters, Wolters, P. J., Adam R. Abate, … Bhattacharya, M. (2019). Reference-based analysis of lung single-cell sequencing reveals a transitional profibrotic macrophage. *Nature Immunology, 20*(2), 163–172. https: //doi.org/10.1038/s41590-018-0276-y

[26] Alexandra Chloé Villani, Villani, A.-C., Rahul Satija, Satija, R., Gary Reynolds, Reynolds, G., Siranush Sarkizova, Sarkizova, S., Sarkizova, S., Karthik Shekhar, Shekhar, K., James Fletcher, Fletcher, J., Morgane Griesbeck, Griesbeck, M., Andrew Butler, Butler, A., Shiwei Zheng, Zheng, S., … Hacohen, N. (2017). Single-cell RNA-seq reveals new types of human blood dendritic cells, monocytes, and progenitors. *Science, 356*(6335). https: //doi.org/10.1126/science.aah4573

[27] Luca Cassetta, Cassetta, L., Stamatina Fragkogianni, Fragkogianni, S., Andrew H. Sims, Sims, A. H., Agnieszka Swierczak, Swierczak, A., Lesley M. Forrester, Forrester, L. M., Hui Zhang, Zhang, H., Daniel Soong, Soong, D. Y. H., Tiziana Cotechini, Tiziana Cotechini, Cotechini, T., Pavana Anur, Anur, P., … Pollard, J. W. (2019). Human Tumor-Associated Macrophage and Monocyte Transcriptional Landscapes Reveal

Cancer-Specific Reprogramming, Biomarkers, and Therapeutic Targets. *Cancer Cell, 35*(4), 588. https: //doi.org/10.1016/j.ccell.2019.02.009

[28] Loems Ziegler-Heitbrock, Ziegler-Heitbrock, L., Petronela Ancuța, Ancuta, P., Suzanne Crowe, Crowe, S. M., Marc Dalod, Dalod, M., Veronika Grau, Grau, V., Derek N.J. Hart, Hart, D. N. J., Pieter J. M. Leenen, Leenen, P. J. M., Yong Jun Liu, Liu, Y.-J., G. Gordon MacPherson, MacPherson, G. G., Gwendalyn J. Randolph, … Lutz, M. B. (2010). Nomenclature of monocytes and dendritic cells in blood. *Blood, 116*(16), 1–20. https: //doi.org/10.1182/blood-2010-02-258558

[29] Christopher J. Halbrook, C. Lyssiotis, M. Magliano, & A. Maitra. (2023). Pancreatic cancer: Advances and challenges. *Cell*. https: //doi.org/10.1016/j.cell.2023.02.014

[30] Ela Elyada, Elyada, E., Mohan Bolisetty, Bolisetty, M., Pasquale Laise, Pasquale Laise, Laise, P., William F. Flynn, Flynn, W. F., Elise T. Courtois, Courtois, E. T., Richard A. Burkhart, Burkhart, R. A., Jonathan A. Teinor, Teinor, J., Pascal Belleau, Pascal Belleau, Belleau, P., Giulia Biffi, … Tuveson, D. A. (2019). Cross-Species Single-Cell Analysis of Pancreatic Ductal Adenocarcinoma Reveals Antigen-Presenting Cancer-Associated Fibroblasts. *Cancer Discovery, 9*(8), 1102–1123. https: //doi.org/10.1158/2159-8290.cd-19-0094

[31] Matthew D. Young, Young, Young, M. D., Sam Behjati, Behjati, S., & Behjati, S. (2018). SoupX removes ambient RNA contamination from droplet based single cell RNA sequencing data. *bioRxiv, 303727*. https: //doi.org/10.1101/303727

[32] Baron M, Baron, M., Adrian Veres, Veres, A., Samuel L. Wolock, Wolock, S. L., Aubrey L. Faust, Faust, A. L., Renaud Gaujoux, Gaujoux, R., Amedeo Vetere, Vetere, A., Jennifer Hyoje Ryu, Ryu, J. H., Bridget K. Wagner, Wagner, B. K., Shai S. Shen-Orr, Shen-Orr, S. S., Allon M. Klein, … Yanai, I. (2016). A Single-Cell Transcriptomic Map of the Human and Mouse Pancreas Reveals Inter- and Intra-cell Population Structure. *Cell Systems, 3*(4), 346–360. https: //doi.org/10.1016/j.cels.2016.08.011

[33] Luke A. J. O'Neill, O'Neill, L. A. J., Rigel J. Kishton, Kishton, R. J., Jeff Rathmell, & Rathmell, J. C. (2016). A guide to immunometabolism for immunologists. *Nature Reviews Immunology, 16*(9), 553–565. https: //doi.org/10.1038/nri.2016.70

[34] Miguel Reina-Campos, Reina-Campos, M., Nicole E. Scharping, Scharping, N. E., Nicole E Scharping, Ananda W. Goldrath, & Goldrath, A. W. (2021). CD8 + T cell metabolism in infection and cancer. *Nature Reviews Immunology*, 1–21. https: //doi.org/10.1038/s41577-021-00537-8

[35] Douglas Hanahan, Hanahan, D., Hanahan, D., Robert A. Weinberg, & Weinberg, R. A. (2011). Hallmarks of cancer: The next generation. *Cell, 144*(5), 646–674. https: //doi.org/10.1016/j.cell.2011.02.013

[36] Zhuo Zhou, Zhou, Z., Zhuo Zhou, Lili Ren, Ren, L., Li Zhang, Zhang, L. Z., Li Zhang, Li Zhang, Jiaxin Zhong, Zhong, J., Xiao Yan, Xiao, Y., Zhilong Jia, Zhilong, J., Li Guo, Guo, L., Jing Yang, Yang, J., … Wang, J. (2020). Heightened Innate Immune Responses in the Respiratory Tract of COVID-19 Patients. *Cell Host & Microbe, 27*(6), 883. https: //doi.org/10.1016/j.chom.2020.04.017

[37] Diether Lambrechts, Lambrechts, D., Els Wauters, Wauters, E., Bram Boeckx, Boeckx, B., Sara Aibar, Aibar, S., Sara Aibar, David Nittner, Nittner, D., Oliver T. Burton, Burton, O. T., Ayse Bassez, Bassez, A., Ayse Bassez, Herbert Decaluwé, Decaluwé, H., Andreas Pircher, … Thienpont, B. (2018). Phenotype molding of stromal cells in the lung tumor microenvironment. *Nature Medicine, 24*(8), 1277–1289. https: //doi.org/10.1038/s41591-018-0096-5

[38] Yunbin Zhang, Zhang, Y., Jingjing Shen, Song, J., Zhongwei Zhao, Zhao, Z., Ming Yang, Yang, M., Ming Chen, Chen, M., Chen, M., Chenglong Liu, Liu, C., Chenglong Liu, Jiansong Ji, Ji, J., Di Zhu, & Zhu, D. (2020). Single-cell transcriptome analysis reveals tumor immune microenvironment heterogenicity and granulocytes enrichment in colorectal cancer liver metastases. *Cancer Letters, 470*, 84–94. https: //doi.org/10.1016/j.canlet.2019.10.016

[39] Rory Stark, Stark, R., Marta Grzelak, Grzelak, M., James Hadfield, & Hadfield, J. (2019). RNA sequencing: The teenage years. *Nature Reviews Genetics, 20*(11), 631–656. https: //doi.org/10.1038/s41576-019-0150-2

[40] Grace Zheng, Zheng, G. X. Y., Jessica M. Terry, Terry, J. M., Phillip Belgrader, Belgrader, P., Belgrader, P., Paul Ryvkin, Ryvkin, P., Zachary Bent, Bent, Z., Ryan J. Wilson, Wilson, R., Solongo B. Ziraldo, Ziraldo, S. B., Tobias D. Wheeler, Wheeler, T. D., Geoff P. McDermott, Geoff McDermott, … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications, 8*(1), 14049–14049. https: //doi.org/10.1038/ncomms14049

[41] Vincent Bernard, Bernard, V., Alexander Semaan, Semaan, A., Jonathan Huang, Jonathan Huang, Huang, J., F Anthony San Lucas, San Lucas, F. A., Feven C. Mulu, Mulu, F. C., Bret M. Stephens, Stephens, B. M., Paola A. Guerrero, Guerrero, P. A., Yanqing Huang, Huang, Y., Jun Zhao, Zhao, J., … Hector A. Alvarez. (2019). Single-Cell Transcriptomics of Pancreatic Cancer Precursors Demonstrates Epithelial and Microenvironmental Heterogeneity as an Early Event in Neoplastic Progression. *Clinical Cancer Research, 25*(7), 2194–2205. https: //doi.org/10.1158/1078-0432.ccr-18-1955