

RESEARCH ARTICLE

Transcriptional and immunohistological assessment of immune infiltration in pancreatic cancer

Brady Bernard¹, Venkatesh Rajamanickam¹, Christopher Dubay¹, Brian Piening¹, Emilio Alonso^{1,2}, Zeljka Jutric^{1,2^{aa}}, Ephraim Tang^{1,2^{ab}}, Pippa Newell^{1,2,3}, Paul Hansen^{1,2,3}, Terry Medler¹, Andrew Gunderson¹, Kristina Young^{1,3}, Carlo Bifulco^{1,4}, Joanna Pucliowska¹, Marka R. Crittenden^{1,3}, Michael J. Gough^{1*}

1 Earle A. Chiles Research Institute, Providence Cancer Institute, Providence Portland Medical Center, Portland, Oregon, United States of America, **2** Liver and Pancreatic Surgery, Providence Cancer Institute, Portland, Oregon, United States of America, **3** The Oregon Clinic, Portland, Oregon, United States of America, **4** Pathology, Providence Portland Medical Center, Portland, Oregon, United States of America

^{aa} Current address: University of California Irvine, Irvine, California, United States of America

^{ab} Current address: London Health Sciences Center, London, Ontario, Canada

* michael.gough@providence.org



OPEN ACCESS

Citation: Bernard B, Rajamanickam V, Dubay C, Piening B, Alonso E, Jutric Z, et al. (2020) Transcriptional and immunohistological assessment of immune infiltration in pancreatic cancer. PLoS ONE 15(8): e0238380. <https://doi.org/10.1371/journal.pone.0238380>

Editor: Francisco X. Real, Centro Nacional de Investigaciones Oncologicas, SPAIN

Received: February 12, 2020

Accepted: August 15, 2020

Published: August 31, 2020

Copyright: © 2020 Bernard et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its Supporting Information files.

Funding: Funding for this manuscript was provided by NIH R01CA182311 (MJG), NIH R01CA244142 (MJG), NIH R01CA208644 (MRC), and the Providence Opportunity Fund. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Pancreatic adenocarcinoma is characterized by a complex tumor environment with a wide diversity of infiltrating stromal and immune cell types that impact the tumor response to conventional treatments. However, even in this poorly responsive tumor the extent of T cell infiltration as determined by quantitative immunohistology is a candidate prognostic factor for patient outcome. As such, even more comprehensive immunophenotyping of the tumor environment, such as immune cell type deconvolution via inference models based on gene expression profiling, holds significant promise. We hypothesized that RNA-Seq can provide a comprehensive alternative to quantitative immunohistology for immunophenotyping pancreatic cancer. We performed RNA-Seq on a prospective cohort of pancreatic tumor specimens and compared multiple approaches for gene expression-based immunophenotyping analysis compared to quantitative immunohistology. Our analyses demonstrated that while gene expression analyses provide additional information on the complexity of the tumor immune environment, they are limited in sensitivity by the low overall immune infiltrate in pancreatic cancer. As an alternative approach, we identified a set of genes that were enriched in highly T cell infiltrated pancreatic tumors, and demonstrate that these can identify patients with improved outcome in a reference population. These data demonstrate that the poor immune infiltrate in pancreatic cancer can present problems for analyses that use gene expression-based tools; however, there remains enormous potential in using these approaches to understand the relationships between diverse patterns of infiltrating cells and their impact on patient treatment outcomes.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests: MJG and MRC have research support from Bristol Myers Squibb, Mavupharma, and Jounce that is not relevant to the subject of this manuscript. Funding for this manuscript was provided by NIH R01CA182311 (MJG), NIH R01CA244142 (MJG), NIH R01CA208644 (MRC), and the Providence Opportunity Fund. The authors declare that these do not represent competing interests relevant to the subject of this manuscript.

Introduction

Pancreatic cancer is commonly characterized by extensive desmoplastic stroma and an environment that is poorly supportive of adaptive immune responses, yet like many other cancers, the degree of T cell infiltrate in pancreatic tumors is correlated with patient outcome [1–4]. T cells in pancreatic tumors face an array of suppressive mechanisms that can limit their ability to control tumors, and it would be beneficial to understand the relationship between T cell infiltration and the presence of other immune populations that positively or negatively regulate immune responses. For this reason, there is significant effort in the field to understand and manipulate the complex immune environment of tumors.

Quantitative immunohistochemistry (IHC) has long represented the gold standard by which tumor infiltrating immune populations can be assessed, and recent advances in multispectral IHC combined with automated image analysis have made possible an unprecedented ability to map out the immune environment of tumors. However, these approaches are limited by the availability and quality of antibodies, and complex multispectral panels require extensive validation to confirm the specificity and selectivity of binding. Recently, multiple groups have shown that the quantity of a diverse array of infiltrating immune cell types in a specimen can be inferred based on characteristic gene expression patterns unique to or enriched in specific cell types [5, 6]. In theory, a single RNA sequencing (RNA-Seq) analysis of preserved tissue can provide an assessment of immune cell infiltration as well as other information such as the cytokine and chemokine balance that may be regulating cell entry and retention in the tissue, together with candidate features of the cancer cells that orchestrate this environment. The addition of simultaneous whole exome sequencing can permit comprehensive profiling of cancer driver mutations, immune targetable mutations, as well as a personalized understanding of the patient's immune profile [7, 8]. However, it remains unclear whether IHC and gene expression-based immune assessment approaches are highly concordant. For example, the utility of RNA-Seq in tumor profiling can be limited by a range of unique factors including degradation of transcripts in excised human tissues and by common tumor preservatives (e.g. formalin) and the ability to detect low-abundant transcripts. The latter is of particular concern in pancreatic cancer, which can have a relatively low infiltration of critical cell types such as CD8 T cells.

Thus far we are not aware of any studies that have directly compared IHC quantification of immune infiltration to RNA-Seq-based analyses in pancreatic cancer. In this study we aim to directly compare conventional IHC and gene expression-based approaches to characterize the immune environment of pancreatic cancer. We hypothesize that RNA-Seq analysis can provide a comprehensive alternative to quantitative IHC for immunophenotyping pancreatic cancer. We performed RNA-Seq on a prospective cohort of 39 pancreatic adenocarcinoma patient tumors with matched quantitative IHC, and evaluated approaches to quantify infiltrating immune cells using gene expression data. We found limited agreement between IHC and RNA-Seq analysis of infiltrating cells, however concordance was greatest when multiple cell types were aggregated to identify a mixed population, such as CD3⁺ T cells from a combination of CD4⁺, CD8⁺, and other cell types that express CD3. This aggregation may overcome the limitation of low T cell-derived RNA transcripts in poorly-immune infiltrated tumors. As an alternative, we identified gene signatures that were enriched in highly T cell infiltrated tumors that are associated with increased disease-free survival in other patient cohorts. These data demonstrate that immune infiltration remains an important predictor of outcome in pancreatic cancer patients, and that RNA analysis can provide an important addition to IHC data to understand the complexities of the immune environment that influence patient outcome.

Methods

Quantitative immunohistology for infiltrating immune cells in pancreatic cancer

We conducted a prospective cohort study of resectable pancreatic masses to determine which immunologic parameters have prognostic value. All procedures were approved under Providence Portland Medical Center Institutional Review Board, with approval number IRB 10–037, and patients provided written informed consent. We restricted our analysis to adult patients who underwent surgical resection for pancreatic masses. Patients were recruited from June 2010 to November 2014 at Providence Portland Medical Center in Portland, OR, where the research was conducted. Inclusion criteria included patients 18 years or older who had a diagnosis of a pancreatic or ampullary mass who were scheduled for surgical resection. Patients had to be able to give informed consent and could not have a diagnosis of a prior malignancy unless they were disease free for 10 years. We included patients who were subsequently determined to have other histologies. Demographics and survival of these patients are outlined in Table 1. Prior studies on this cohort had identified a positive correlation between CD3+ T cell infiltrate and overall survival by Multivariate Cox modeling and univariate analysis [1], so this sample set was applied for additional genomic analysis. Additional power calculations were not performed. Tumor infiltrating immune cells were quantified by immunohistochemistry and quantitative digital image analysis for CD3⁺, CD68⁺, and CD8⁺ cells as previously described [1]. Infiltrating cells were quantified from whole slide digital

Table 1. Demographics of patients on the study.

<i>Characteristic</i>	<i>Data</i>
Pancreatic Adenocarcinoma, n	75
<i>Age, median (range), y</i>	66 (25–87)
<i>Male, n (%)</i>	43 (57%)
<i>Female, n (%)</i>	32 (43%)
Neoadjuvant treated PDA, n	6
<i>Age, median (range), y</i>	61 (58–61)
<i>Male, n (%)</i>	3 (50%)
<i>Female, n (%)</i>	3 (50%)
Premalignant, n	13
<i>Age, median (range), y</i>	58 (28–79)
<i>Male, n (%)</i>	5 (38%)
<i>Female, n (%)</i>	8 (62%)
Benign, n	10
<i>Age, median (range), y</i>	53 (40–71)
<i>Male, n (%)</i>	5 (50%)
<i>Female, n (%)</i>	5 (50%)
Neuroendocrine, n	3
<i>Age, median (range), y</i>	61 (55–69)
<i>Male, n (%)</i>	2 (66%)
<i>Female, n (%)</i>	1 (33%)
Duodenal Adenocarcinoma, n	1
<i>Age, median (range), y</i>	61
<i>Male, n (%)</i>	1 (100%)
<i>Female, n (%)</i>	0 (0%)

<https://doi.org/10.1371/journal.pone.0238380.t001>

images scanned at 20x resolution (Leica SCN400). Regions of interest were defined with Pathologist guidance using Definiens Tissue Studio (Definiens Inc), and the automated algorithm used the immunohistology staining combined with nuclear counterstain to count total cells and positive cells to report a marker positive cell density per mm² tissue for each patient. Primary outcome was overall survival.

RNASeq analysis of pancreatic cancer

Of the patients above with pancreatic masses, we randomly selected 39 patients with pathologically diagnosed pancreatic adenocarcinoma with no neoadjuvant treatment and with matched quantitative IHC for RNASeq. All subsequent analyses of IHC and RNASeq were performed on untreated patients. A representative Hematoxylin and Eosin (H&E) stained slide for each formalin-fixed paraffin-embedded (FFPE) tissue block specimen was reviewed by a board-certified pathologist for tumor content and tumor-rich regions were identified for microdissection. Blocks were matched but not in series with IHC sections. 5 μm thick unstained sections on glass slides were processed for DNA and RNA purification by the Providence Molecular Genomics Laboratory. The FFPE tissue sections were deparaffinized using Envirene (Hardy Diagnostics) followed by RNA extraction and purification using the Qiagen AllPrep DNA/RNA FFPE kit. 85ng of input RNA was used to prepare sequencing libraries using the Illumina TruSeq RNA Exome kit. Sequencing of the RNA Exome libraries was performed on the Illumina HiSeq 4000 instrument at 2 x 75 read paired end configuration. Transcripts were quantified using salmon-v.0.11.2 [9]. A matrix of gene expression values for all patients analyzed in this study along with matched quantitative IHC are provided as a [S1 Table](#).

Computational analysis of infiltrating cells and comparison of techniques

RNA-Seq-based cell type deconvolution was performed using xCELL (5) and CIBERSORT (6), with TPM gene expression levels as input. xCELL was used to perform *cell type enrichment analysis* from gene expression data for 64 immune and stroma cell types, whereas CIBERSORT provides *absolute and relative abundance* of different immune cell types depending on the specified gene set. In the present analysis, the LM22 signature provided by CIBERSORT was applied. We also applied EPIC (10) and MCPcounter (11) to estimate the abundance of immune cell population. EPIC estimates the proportions of Immune and Cancer cells by using RNA-Seq-based gene expression reference profiles from immune cells and other nonmalignant cell types found in tumors. MCPcounter quantifies abundance of tissue-infiltrating eight immune and two stromal cell populations based on transcriptome profile.

Initial clustering of patients based on infiltrating cell types was performed using ClusterVis [10]. Principal components are calculated as described in (6) using ClusterVis (7). Missing data is assigned using Singular Value Decomposition with imputation iteratively until estimates of missing values converge. Statistical significance of the resulting hierarchical clusters were assessed using the sigclust2 R package [11]. Correlation between infiltrating cell types calculated using RNA-Seq versus quantitative IHC was performed by generating a composite of T cell and macrophage cell types determined by RNA-Seq for direct comparison to IHC populations (Table 2). Correlations between xCELL immune cell type enrichment and quantitative IHC, as well as between CIBERSORT cell type abundances and quantitative IHC, were determined by Spearman Rank correlation.

Identification of genes associated with high and low T cell infiltration

Patients were categorized into high and low infiltration of CD3⁺ and CD8⁺ T cells according to quantitative IHC and sub-categorized into those with high infiltration of both populations

Table 2. Assembly of composite RNASeq-calculated populations to match cell types detected by IHC.

IHC populations	CD3		CD8	CD68
CIBERSORT	T.cells.CD8	T.cells.CD4.naive	T.cells.CD8	Macrophages.M0
	T.cells.CD4.memory.resting			Macrophages.M1
	T.cells.CD4.memory.activated	T.cells.follicular.helper		Macrophages.M2
	T.cells.regulatory.Tregs	T.cells.gamma.delta		
xCELL	CD4+ memory T-cells	CD4+ naive T-cells	CD8+ naive T-cells CD8+ T-cells CD8+ Tcm CD8+ Tem	Macrophages
	CD4+ T-cells	CD4+ Tcm		Macrophages M1
	CD4+ Tem	CD8+ naive T-cells		Macrophages M2
	CD8+ T-cells	CD8+ Tcm		
	CD8+ Tem	Tregs		
	Th1 cells	Th2 cells		
	Tgd cells	NKT		

<https://doi.org/10.1371/journal.pone.0238380.t002>

(CD3^{HI}CD8^{HI}) versus low infiltration of both populations (CD3^{LO}CD8^{LO}). Using RNA-Seq data from these patients, gene expression analyses were performed using a univariate two-sample T-test with a stringent false-positive threshold to identify genes significantly differentially expressed ($p < 0.001$) in our patients [12, 13]. These genes were mapped to known pathways using the Reactome Functional Interaction network tool [14]. These genes were tested on the TCGA Pancreatic Adenocarcinoma PanCancer Atlas [15] as a validation cohort using cBioportal [16, 17] where mRNA expression z-scores are compared to the expression distribution of each gene in tumors that are diploid for this gene. To be classified as enriched for the gene score the sample must have at least one gene that is 2 log fold overexpressed. Survival and expression data exported to Graphpad Prism for survival comparison using log-rank tests. Pre-calculated xCELL analysis of patients in the TCGA database were obtained from <http://xcell.ucsf.edu>.

Statistical methods

Survival data were analyzed using Prism (Version 8.4.2, GraphPad Software, La Jolla, CA). Overall survival of groups was compared using a log rank test for differences in Kaplan-Meier survival curves. All cutoffs for high/low infiltration by RNA analysis use median values. Gene expression analyses were performed using a univariate two-sample T-test with a stringent false-positive threshold to identify genes significantly differentially expressed ($p < 0.001$) [12, 13]. Correlations between immune cell type enrichment and quantitative IHC were determined by Spearman Rank correlation. To assess the statistical significance of correlation between RNA-Seq-based cell type deconvolution and CD3, CD8 or CD68 immune cell types determined by quantitative IHC, the composition of underlying cell types specified in Table 2 were randomly permuted. The number of cells that constitute a cell type were kept the same between the true set and permuted set. Owing to the number of cell types estimated by xCELL (64 different cell type populations) and CIBERSORT (22 different cell type populations), we conducted 100 and 20 different random assignments of cell types, respectively, attributed to the CD3, CD8 and CD68 IHC populations in Table 2. For each permutation, Spearman rank correlation was computed between the random cell type assignments and quantitative IHC, allowing for a level of statistical significance to be estimated for the true set with respect to all permuted sets, thereby indicating whether the RNA-Seq-based cell type deconvolution methods are statistically significantly correlated with quantitative IHC. Additional correlations between multiple variables were analyzed using Prism (Version 8.4.2) to calculate a Pearson correlation coefficient. Statistical significance of hierarchical clusters were assessed using the

sigclust2 R package [11]. Differential gene expression analysis between CD3^{Hi} CD8^{Hi} vs CD3^{Lo} CD8^{Lo} RNA-Seq samples were carried out using DESeq2 [18]. We identified differentially expressed genes, ranked them based on fold-change and p-value. This gene signature was further used for gene set enrichment.

Results

We previously demonstrated that increased CD3⁺ T cell infiltrates in surgically resected pancreatic adenocarcinoma correlate with improved outcomes using a Cox proportional hazards model, and remained prognostic by multivariable analysis [1]. By contrast, CD8⁺ T cell infiltrates and CD68⁺ macrophage infiltrates did not correlate with outcome [1]. However, using categorical variables of high or low CD3⁺, CD8⁺ and CD68⁺ cell numbers in tumors, we were able to identify cutoffs that demonstrated that patients with high CD3⁺ T cells or high CD8⁺ T cells exhibited improved survival, but again CD68⁺ macrophages did not correlate with survival (Fig 1a). To evaluate the complexity of the infiltrate in patient tumors, we performed hierarchical cluster analysis. Initially, we used our broader dataset that included pancreatic ductal adenocarcinoma (PDA) with and without neoadjuvant treatment, as well as benign pancreatic masses, pre-malignant disease, neuroendocrine tumors, as well as a small number of duodenal adenocarcinoma and gallbladder adenocarcinoma. Principal component analysis was not able to distinguish these pathologies based on CD3⁺, CD8⁺ and CD68⁺ cell infiltrate (Fig 1bi), and while cluster analysis tended to group benign and premalignant disease in poorly infiltrated groups, there was not a clear classifier to distinguish PDA from other related pathologies (Fig 1bii). Prior studies have demonstrated that patients with the highest macrophage proportions and lowest CD8⁺ T cell proportions exhibit worse outcome than those with lowest macrophage proportions and highest T cell proportions [19]. We calculated the correlation between CD3⁺, CD8⁺ and CD68⁺ cell infiltrate in patients with all pathologies and those with PDA and found good correlation between CD3⁺ and CD8⁺ infiltrates, and poor, but not negative, correlation with T cells and CD68⁺ cell infiltrate (Fig 1c). To determine whether the degree of CD68⁺ cell infiltrate impacted outcome for patients with high or low T cell infiltrates, we tested the effect of CD68⁺ cell infiltrate with each T cell cutoff. For patients with PDA we did not find an effect of macrophage infiltration on the survival benefit of CD3⁺ T cells (Fig 1di and 1dii) or CD8⁺ T cells (Fig 1diii and 1div). These data demonstrate that quantitative immunohistology was able to identify good and poor outcome groups based on CD3⁺ and CD8⁺ T cell infiltrate, but analyzing the degree of CD68⁺ macrophage infiltrate alone or in combination with T cell infiltrates did not help refine outcome groups.

There are significant limitations in the use of CD68 as a sole marker of macrophages in tumors [20], particularly in view of the diverging phenotypes macrophages can generate. There is not a well-defined set of markers that is unique to distinct macrophage polarization states that do not overlap with other cell types. Recently, a number of algorithms have been developed that can analyze gene expression data to estimate the prevalence of the broad range of cell types in a mixed tissue sample [5, 6, 21]. To evaluate whether gene expression analysis could refine our understanding of the immune environment of pancreatic cancer we tested two different approaches, CIBERSORT [6] and xCELL [5]. CIBERSORT has been most widely applied, and provides an assessment of 22 of the most common immune cell types and some information on the differentiation of CD4⁺ T cells and macrophages [6]. We performed RNA-Seq on a subset of our PDA patients with quantitative IHC, and performed CIBERSORT analysis of immune infiltration using the RNA-Seq data. We used a correlation analysis to determine whether some cell types were co-regulated in the tumor, but there was little evidence of correlation between the infiltration of different immune cell types in the tumor

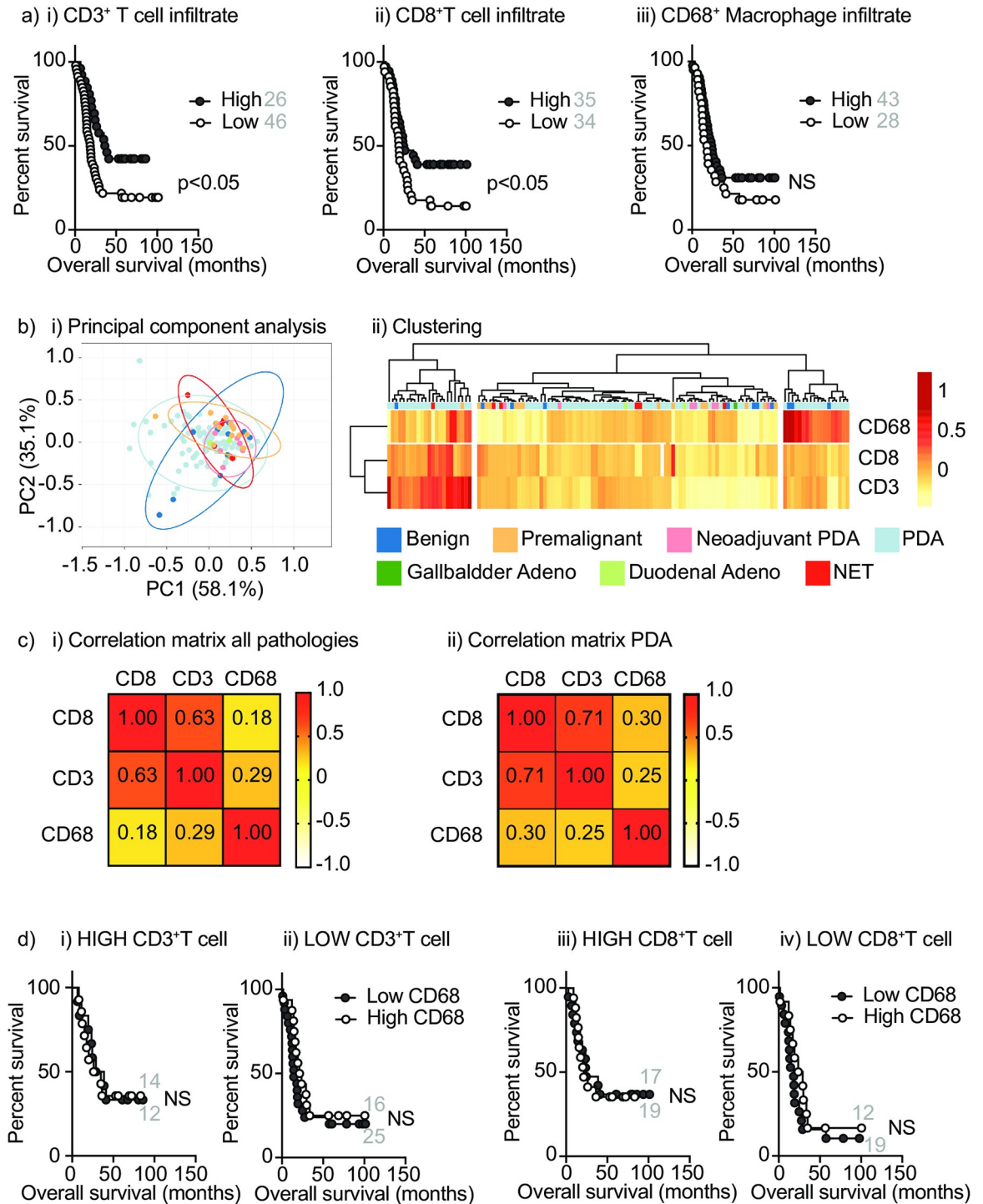


Fig 1. Analysis of patients by quantitative immunohistochemistry. a) Quantitative analysis of infiltrating i) CD3⁺; ii) CD8⁺; and iii) CD68⁺ cells/mm² tumor were used to set cutoff values of high/low infiltrates with significant impact on overall survival of patients with PDA. Cutoffs used were approximately CD3 – 75th percentile; CD8 – median; CD68 – no significant cutoff found, 75th percentile shown. b) Infiltrating CD8⁺, CD3⁺, and CD68⁺ cells across a range of pathologies was used to evaluate principal component analysis and clustering. i) Unit variance scaling is applied to rows; SVD with imputation is used to calculate principal components. X and Y axis show principal component 1 and principal component 2 that explain 56.7% and 30.5% of the total variance, respectively. Prediction ellipses are such that with probability 0.95, a new observation from the same group will fall inside the ellipse. N = 123 data points. ii) Clustering of patients according to infiltrates. Imputation is used for missing value estimation. Both rows and columns are clustered using Manhattan distance and complete linkage. 3 rows, 123 columns. c) Pearson correlation coefficients for CD8, CD3 and CD68 infiltrating cells in i) all

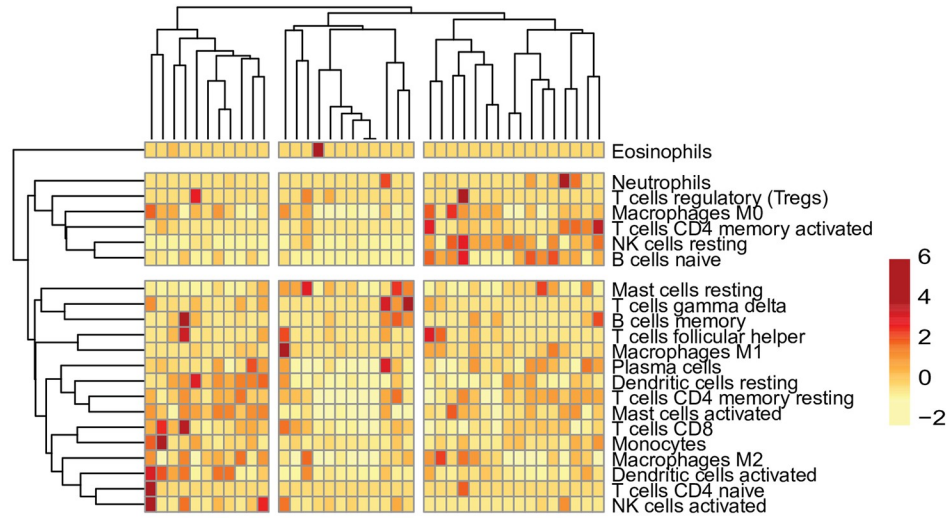
pathologies; ii) PDA. d) Overall survival of PDA patients with i) high or ii) low CD3⁺ infiltrates, and iii) high or iv) low CD8⁺ infiltrates subdivided according to high or low CD68 infiltrates as determined in a). Number of patients on each arm of survival curves are shown in grey.

<https://doi.org/10.1371/journal.pone.0238380.g001>

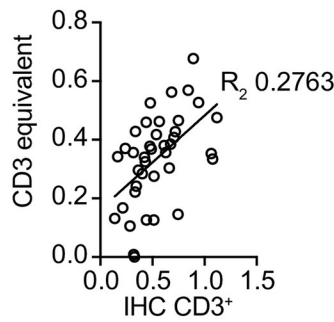
(S1 Fig). We performed cluster analysis on patients based on their immune infiltrates calculated by CIBERSORT (Fig 2ai), which appeared to identify a diffuse cluster of patients with higher numbers of CD8 T cells and dendritic cells, but there was no strong statistical association between these cell types and this analysis was not able to identify patient groupings with significant differences in overall survival (not shown). To directly compare these CIBERSORT calculated infiltrating cell types to the quantified immunohistology from the same samples, we made three combined groups (Table 2): 1. CD3⁺ equivalent based on cell types that express CD3 (Treg+CD4 populations+CD8+gd T cells); CD8⁺ equivalent (CD8); and CD68⁺ macrophage equivalent (MO+M1+M2). We then evaluated the correlation between infiltration determined by CIBERSORT analysis of RNA-Seq versus quantified immunohistology. Both CD3⁺ and CD8⁺ T cells showed a weakly positive correlation between the IHC and CIBERSORT assessments, but CD68 did not correlate well between histology and CIBERSORT (Fig 2b). Spearman's rho computed between IHC and CIBERSORT relative cell type abundance were 0.355, 0.308, 0.144 for CD3, CD8 and CD68 respectively. Notably, we see a number of patients with no detectable infiltrating T cells or macrophages by CIBERSORT, who did have detectable cells by histology. Since CIBERSORT is dependent on key RNA transcripts being present amongst the RNA sequenced, we believe that at low cell infiltration this approach can struggle to identify the RNA signature of rare infiltrating cells. To determine whether CIBERSORT infiltration of these key cell types predicted outcome, we similarly stratified patients into high or low infiltration groups. We found that patients with high combined CD3⁺ equivalent scores exhibited improved overall survival ($p < 0.05$), but the CD8 and macrophage scores were not able to discriminate patients with significantly different overall survival (Fig 2c). These suggest that while CIBERSORT analysis can provide additional information on the diversity of immune cells in tumors, it does not improve our ability to predict outcome over quantitative histology. Since there is a general agreement between the assessment of total CD3 infiltrate by histology and CIBERSORT and each are associated with improved outcome in patients, aggregating CIBERSORT T cell infiltration could be further tested as a prognostic factor in pancreatic cancer patients.

As an alternative approach xCELL can identify 64 different cell populations and composite infiltration scores from RNA-Seq data [5]. We performed xCELL analysis of immune infiltration in pancreatic cancer and generated a correlation matrix to examine associations between different cell types. Clear patterns emerged, with some tight clusters based around epithelial cells or Th2 cells, and more broad groupings of co-regulated cells including Th1 cells, DC, M1 macrophages and CD8 T cells (S2 Fig). To determine whether these co-resident cells identified unique patient populations, we clustered patients based on their immune infiltrate, identifying patients with higher levels of CD8 T cells, DC, and M1 macrophages (Cluster A), and those with higher levels of fibroblasts and endothelial cells to form a distinct cluster (Cluster B) (Fig 3a). Comparing the overall survival of patients in each cluster demonstrated there were no significant differences between groups (Fig 3b). In view of the high correlation between specific cell types resulting in apparent clusters, we investigated whether this was due to closely related cells having overlapping genes that are included in the underlying gene signature. Using the gene signatures that determine cell types in xCELL, we computed the percent match between genes across all cell types (S3 Fig). These data demonstrated that most cells were defined using a unique gene set. We did identify overlapping gene usage between related cell types such as

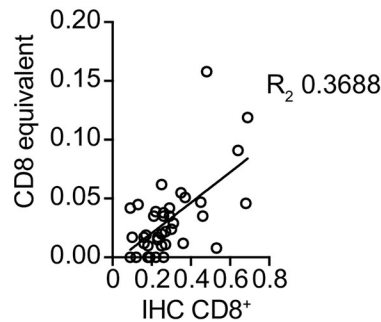
a) Clustering PDA patients based on CIBERSORT analysis of immune infiltration



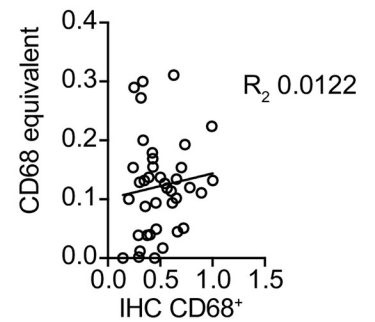
b) i) CD3 correlation



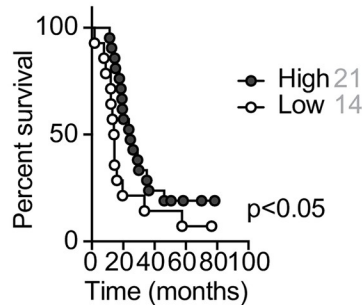
ii) CD8 correlation



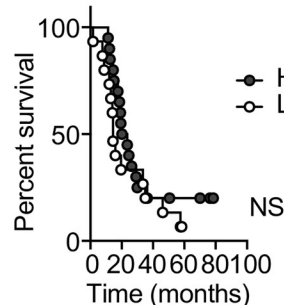
iii) CD68 correlation



c) i) CD3 equivalent



ii) CD8 equivalent



iii) CD68 equivalent

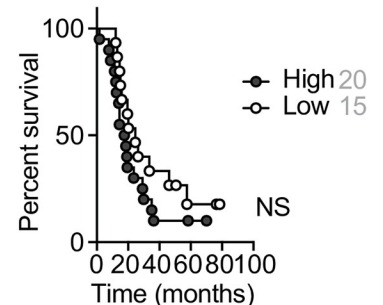


Fig 2. Evaluation of CIBERSORT analysis of immune infiltration. a) Clustering of immune populations inferred using CIBERSORT analysis of RNASeq of PDA patients. Rows are centered; unit variance scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage. 22 rows, 39 columns. b) The immune infiltrate score of all i) CD3⁺, ii) CD8⁺, and iii) CD68⁺ equivalent cell populations (Table 2) determined by CIBERSORT compared to quantitative IHC from the same patient. Each symbol represents one patient. c) Overall survival of patients with high versus low infiltrates of i) CD3⁺, ii) CD8⁺, and iii) CD68⁺ equivalent cell populations (Table 2) determined by median CIBERSORT infiltration. NS = not significant.

<https://doi.org/10.1371/journal.pone.0238380.g002>

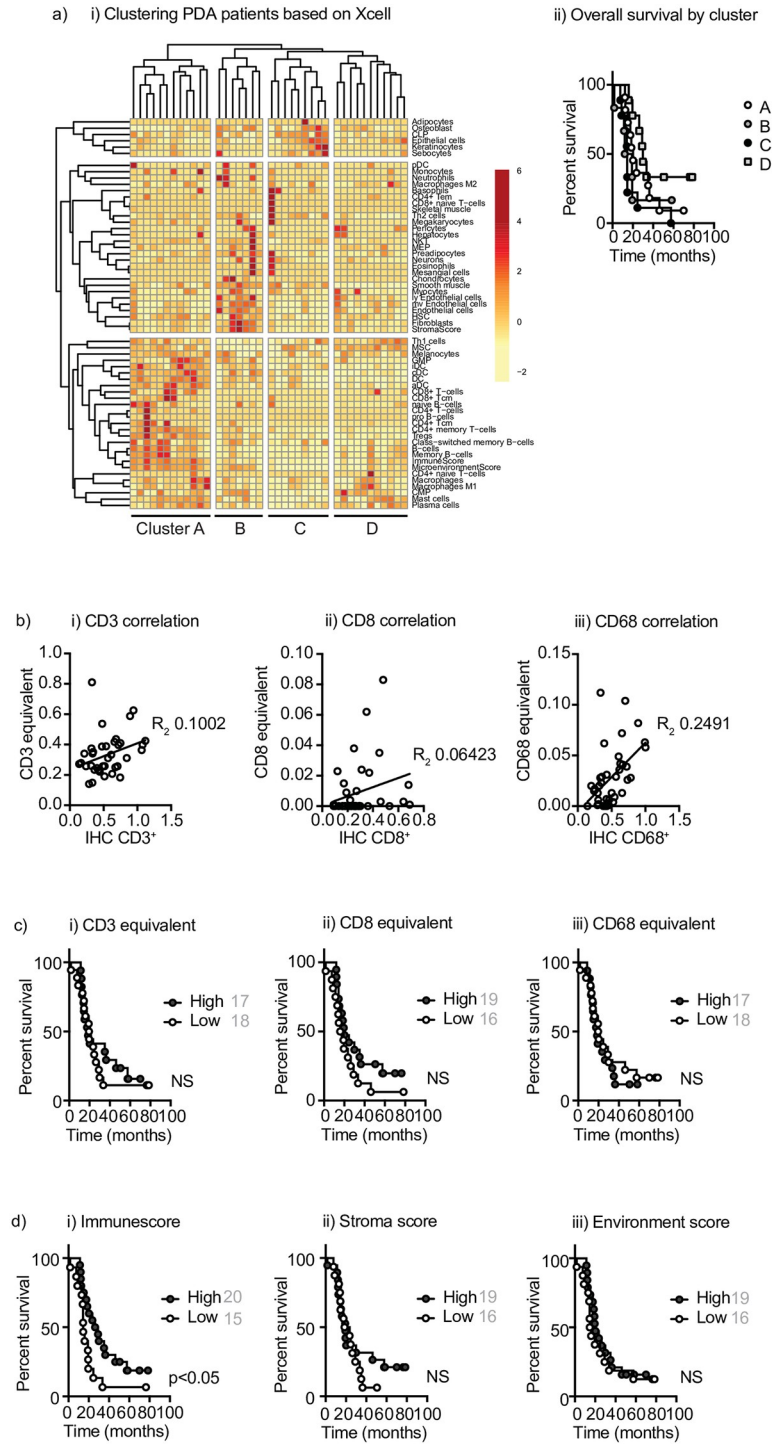


Fig 3. Evaluation of xCELL analysis of immune infiltration. a) i) Clustering of immune populations inferred using xCELL analysis of RNASeq of PDA patients. Rows are centered; unit variance scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage. 60 rows, 39 columns. ii) overall survival of patients in clusters A, B, C, and D. b) The immune infiltrate score of all i) CD3⁺, ii) CD8⁺, and iii) CD68⁺ equivalent cell populations (Table 2) determined by xCELL compared to quantitative IHC from the same patient. Each symbol represents one patient. c) Overall survival of patients with high versus low infiltrates of i) CD3⁺, ii) CD8⁺, and iii) CD68⁺ equivalent cell populations (Table 2) determined by median xCELL infiltration. NS = not significant. d) Overall survival of patients with high versus low i) immunescore, ii) stromascore, and iii) environment score using median values as cutoffs.

<https://doi.org/10.1371/journal.pone.0238380.g003>

DC subtypes or T cell subtypes; however, there were no significant gene overlaps between DC and T cells, for example, that would explain their correlation. These suggest that the positive correlation between the number of T cells and DCs in patient tumors is likely a result of the presence of both cell types in the analyzed sample. To compare these calculated infiltrating cell types to quantified immunohistology, we again made three combined groups: 1. CD3 equivalent (Treg + all CD4 populations + all CD8 populations + gd T cells); CD8 equivalent (all CD8 populations); and CD68 macrophage equivalent (MO + M1 + M2) (Table 2). We evaluated the correlation between cell infiltration assayed by xCELL analysis of RNA-Seq versus quantified immunohistology. Each population showed a positive correlation between the two approaches; however, many of the samples were calculated to have no CD8 T cells by xCELL, even in patients with relatively abundant CD8 T cells as determined by immunohistology (Fig 3b). As with CIBERSORT analysis, despite the positive correlation, the R^2 value was not strong for each cell type. Spearman's rho computed between IHC and xCELL were 0.354, 0.307 and 0.144 for CD3, CD8 and CD68 respectively. We also evaluated whether infiltration of these cell types impacted outcome, and we could not detect a cutoff that impacted overall survival (Fig 3c).

xCELL analysis does calculate three additional fields that integrate many of the infiltrating cell features to generate an "immunescore", a "microenvironment score" and a "stromascore". Such combined fields may have an advantage over individual cell types, particularly where the cells are of low abundance. Broadly, the immunescore was correlated with T cell infiltration, while the stromascore was correlated with fibroblast and endothelial cell infiltration (S3 Fig). Patients with a high immunescore exhibited improved overall survival, but the stromascore and the microenvironment score were not able to distinguish patient groups with improved outcome. These data suggest that xCELL can provide a more complex understanding of the immune cell diversity in tumors; however, there remain significant issues identifying the small numbers of T cells infiltrating pancreatic adenocarcinoma. There may be a benefit in integrating multiple immune cell types through features such as the immunescore to identify tumor environments indicative of improved outcome.

There are increasing numbers of methods to analyze cell infiltrates from RNA data. We compared the additional methods EPIC [22] and MCPcounter [23] with the same dataset. The MCPcounter assessment of T cell infiltration correlated well with the IHC CD3 infiltrate, but all other analyses had poor correlation to IHC data (S4 Fig). To understand whether there was agreement amongst the various RNA-based approaches, we analyzed the correlation between the various infiltrating T cell populations assessed by CIBERSORT, xCELL, MCPcounter, and EPIC. The correlation between the different approaches using the same RNA dataset was moderate, but the closest correlations were found among CD4 T cell populations and relatively poor correlations between CD8 T cell populations (S4 Fig). These data suggest that there are significant differences between the RNA-based approaches and each has difficulty consistently identifying CD8 T cell infiltrates in T cell poor tumors like pancreatic cancer.

To determine whether there is an alternative RNA signature of high T cell infiltration that can be used in pancreatic tumors to infer T cell infiltration and assess outcome using RNA-Seq samples, we identified genes that were enriched in tumors with both high CD3 and high CD8 infiltration or both low CD3 and low CD8 infiltration by quantitative IHC. Class comparison of gene expression identified a subset of genes that were statistically associated with highly T cell infiltrated tumors (Table 3, Fig 4a). As would be predicted, these included genes encoding for CD3 as well as genes involved in T cell signaling such as FYN and LAT. Interestingly, the gene set includes the immunotherapy target CTLA4 [24], as well as SLAMF6, which is a marker of progenitor exhausted T cells [25]. To determine whether increased expression of these genes were useful predictors of outcome in pancreatic cancer patients, we examined expression of these genes in pancreatic adenocarcinoma patients in the TCGA database [15],

Table 3. Normalized log transformed expression of genes that are statistically associated with highly T cell infiltrated tumors.

Class	HI	HI	HI	HI	HI	HI	HI	HI	HI	HI	HI	LOW	LOW	LOW	LOW	LOW	LOW	LOW	LOW	LOW	LOW	Ttest p-value	DISEQ2 p-value								
HZR	3309	3101	4975	4716	3512	4606	3344	5294	6196	1856	2738	4568	6144	2479	4679	1489	3159	1141	917	1818	1915	2271	1808	2504	1717	1937	541	1007	3.0E-06	0.00001	
TESPA1	124	181	272	200	182	275	185	158	131	178	234	123	239	85	143	056	189	014	025	076	000	092	000	000	000	000	056	158	025	5.0E-06	0.01095
SLAMF1	171	166	186	153	194	205	186	200	187	199	194	206	227	117	148	053	188	064	047	163	117	124	000	000	160	109	017	023	9.0E-06	0.03359	
KATNAL2	116	110	080	110	054	159	053	133	174	149	094	127	089	064	077	022	070	070	035	017	028	064	000	000	072	037	058	023	9.9E-06	0.17789	
BSP2	226	455	380	358	328	770	561	366	481	191	248	417	353	181	381	169	307	093	027	147	203	114	128	000	140	241	069	021	2.0E-05	0.87588	
MIAMF6	230	312	474	262	311	382	435	344	608	447	204	506	447	212	515	114	388	033	072	235	237	287	000	178	113	236	067	130	2.4E-05	0.00020	
MCOL2	215	212	152	245	206	415	459	189	311	313	220	249	444	221	233	135	200	026	048	097	221	084	050	042	083	095	042	042	3.27E-05	0.00000	
PKRCE	069	092	091	125	131	116	037	127	225	104	047	123	119	108	211	048	068	026	078	048	056	014	000	065	032	032	013	086	0.000119	0.00003	
RDHS	277	288	416	364	369	245	533	309	1145	473	253	577	509	268	350	001	245	002	001	258	003	477	000	000	184	052	000	001	0.000159	0.00695	
GACN4H1	012	044	075	043	055	068	042	058	071	032	000	058	099	022	045	014	007	009	003	019	000	043	044	000	000	000	000	003	0.000179	0.000232	
FAMPC	1291	059	915	1156	1941	364	1179	1941	1184	969	443	1849	351	533	2241	516	2021	2698	2842	407	2948	4283	3446	3729	1178	3321	2553	3185	0.000008	0.00000	
RASGRP2	743	1049	1213	882	939	1243	848	1229	1181	922	461	1041	1491	399	2165	445	813	094	361	881	437	1078	000	000	499	611	089	192	0.000243	0.000015	
PVALB	082	135	198	098	067	084	025	118	150	048	065	070	127	029	117	035	063	013	011	025	039	000	000	000	038	000	038	013	021	0.000239	0.000138
CD3G	1232	893	1244	1073	1499	1395	965	1006	652	1396	625	1348	2085	1203	1486	468	1184	287	296	693	867	980	034	961	985	797	190	381	0.000299	0.002149	
CD3D	5587	2713	4082	5887	3877	7328	4279	4260	2033	6971	2477	4864	8630	3596	6484	1799	5076	799	1286	2656	4104	4076	000	000	2982	2072	306	1431	0.000143	0.01063	
DNND1C	597	488	482	538	690	800	447	403	294	638	538	553	793	405	645	276	523	259	183	282	421	573	000	000	412	519	022	283	0.000418	0.000066	
CD3E	4056	4242	5630	4507	6307	5670	3561	3914	4075	4451	2208	4456	8882	3533	6143	1832	4881	1452	1104	1937	3577	3425	1635	4624	2972	2981	568	1246	0.000308	0.000016	
ALDH	784	254	864	357	902	1151	607	773	287	980	654	185	1434	812	893	197	588	187	157	462	322	383	000	000	558	492	169	204	0.000128	0.007145	
TSPAN7	1492	2409	1161	1469	1414	1160	633	1384	2089	1025	560	1287	1243	592	2652	801	885	427	315	963	312	719	000	521	1340	705	593	191	0.000495	0.000764	
KLRB1	1077	2609	2384	1738	1870	1656	1888	2295	1547	2001	1600	2441	2587	1944	3276	1252	1352	609	901	1274	1286	1624	2690	000	1454	933	359	689	0.000589	0.000001	
ZC3H2D	032	041	041	035	053	078	043	044	048	065	035	040	045	037	047	031	059	026	016	039	035	034	000	013	030	019	008	015	0.000571	0.0001851	
FTN	1068	905	998	894	1526	2090	1146	1368	1113	911	895	1038	1612	910	1327	415	1156	514	629	1116	1107	960	000	704	889	204	497	0.000542	0.063184		
LAT	1319	275	173	143	248	209	371	286	252	401	137	183	293	163	334	105	181	119	097	153	216	207	197	033	138	114	023	051	0.000581	0.683191	
NEBL	348	163	390	255	335	250	215	265	178	347	224	416	673	176	311	347	587	601	556	616	867	412	1211	966	375	545	267	1352	0.000584	0.000026	
HSP90A1	22153	35300	25865	24609	30787	26880	26816	20139	002	15303	99519	26095	22847	16512	19869	14989	27635	48185	55154	37817	43315	49492	48183	45889	36547	48720	16487	46971	0.000602	0.028536	
RAB9B	023	019	020	034	015	028	005	026	031	030	016	065	018	020	029	014	022	009	002	010	018	015	000	000	010	003	010	008	0.000729	0.001085	
TMEM156	294	466	370	358	542	596	342	431	000	249	154	545	760	393	600	173	427	256	183	200	193	045	000	086	351	117	043	167	0.000765	0.026697	
MCMB	021	908	247	487	289	368	492	072	711	236	030	506	633	131	426	000	332	107	000	000	309	134	000	000	037	016	015	000	0.000729	0.807953	
CTLA4	426	521	641	536	910	1090	329	517	414	444	408	506	1016	495	481	213	882	169	069	235	403	304	000	000	365	212	053	305	0.000741	0.000770	
CSTF	30666	21933	17257	37521	24166	25596	14999	21406	11222	15454	27536	21492	17351	18579	25847	19488	39389	31594	42305	48419	61207	38577	55417	78562	37477	46959	6051	0.000762	0.010199		
PYHIN1	296	255	418	225	212	740	259	289	410	327	372	376	826	248	653	187	330	022	106	312	186	279	000	000	163	176	093	241	0.000790	0.008206	
SIA	1743	054	1534	1703	1580	130	1168	1786	1916	1244	119	1458	2030	1212	1801	1244	155	098	000	151	1192	143	1816	000	099	000	012	036	0.000836	0.170331	
CD63	513	422	427	353	671	623	404	333	000	566	000	355	638	467	476	000	516	000	000	000	395	361	000	370	000	000	000	065	085	0.000948	0.128691
SHISA2A	104	071	141	175	223	167	084	094	231	144	089	159	144	065	137	061	123	040	020	046	078	165	000	000	051	157	011	016	0.000950	0.000681	
PPKCD1	139	1822	136	164	237	250	1593	239	133	198	2023	124	314	208	136	1204	300	1549	1576	179	2584	2787	3103	1629	1935	603	1979	0.000976	0.57628	0.000954	0.001094
PPP1R1B	204	234	186	131	384	392	173	247	279	284	176	236	390	182	256	102	273	079	104	156	175	178	000	170	221	206	024	082	0.000954	0.001094	

<https://doi.org/10.1371/journal.pone.0238380.t003>

p=0.2

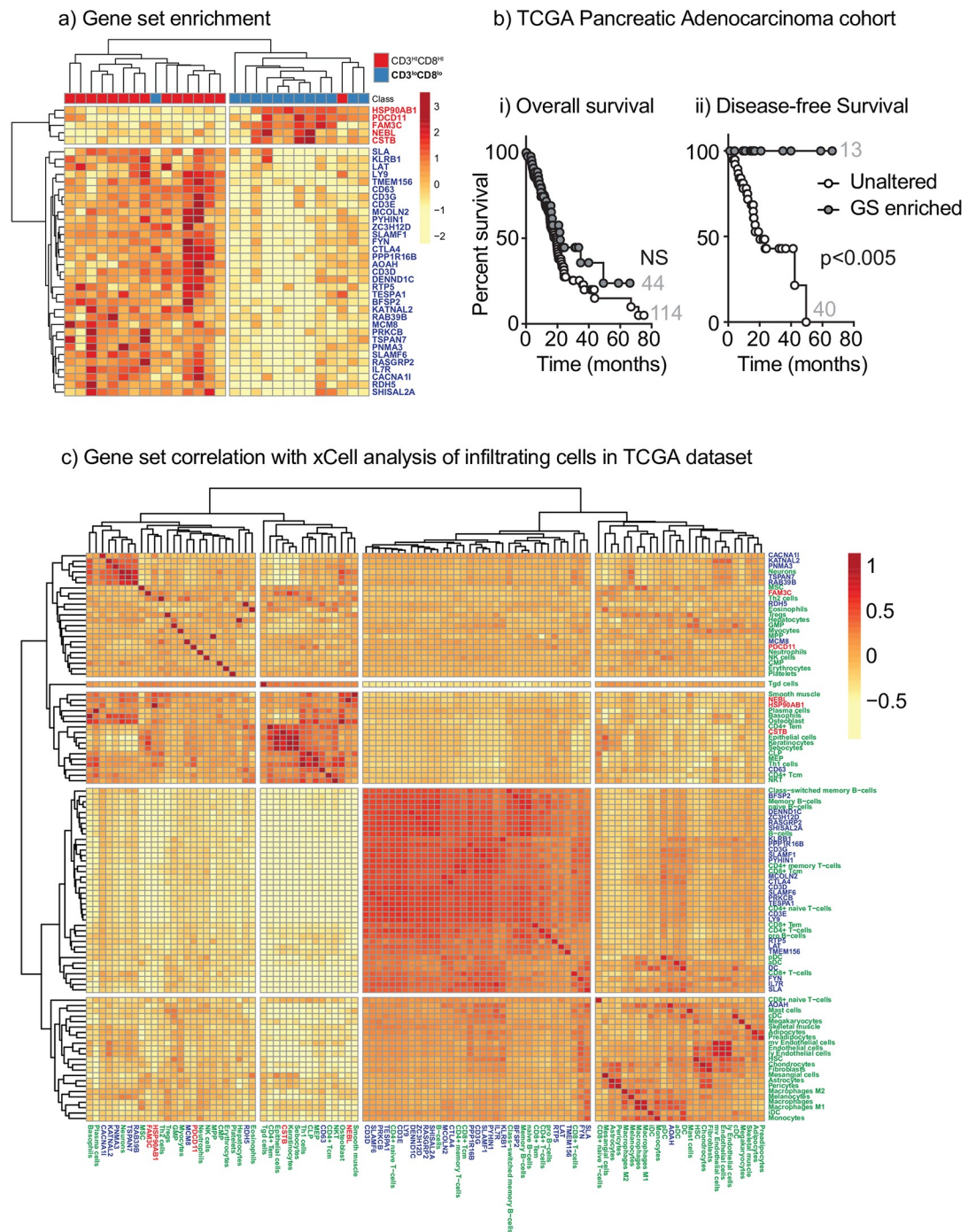


Fig 4. Class comparison of TPM counts on patients with high T cell infiltrate and validation on TCGA cohort. a) Class comparison of TPM counts was performed on patients with high CD3⁺ and high CD8⁺ infiltrates, as compared to patients with low CD3⁺ and low CD8⁺ infiltrates using the cutoffs applied in Fig 1. Graph shows the clustering of 38 significant genes (p < 0.001). Genes enriched in poorly infiltrated tumors are highlighted in bold. b) The genes enriched in highly infiltrated tumors were tested using the TCGA PanCan Pancreatic Adenocarcinoma cohort to identify patients with enriched RNA expression of these genes. Graphs show i) overall survival and ii) disease-free survival for patients in each group. The number of patients in each group is shown. c) Correlation matrix of the expression of the genes listed in a), and xCELL calculated cell infiltration for the TCGA PanCan Pancreatic Adenocarcinoma cohort. Rows are centered; no scaling is applied to rows. Both rows and columns are clustered using Manhattan distance and average linkage. 102 rows, 102 columns. Genes are highlighted in Blue where positively correlated and Red when negatively correlated. Cell types are highlighted in Green.

<https://doi.org/10.1371/journal.pone.0238380.g004>

and their effect on patient outcome. Initial analysis indicated that patients with increased expression of the genes positively associated with T cell infiltration in our cohort had significantly increased overall survival and disease-free survival. However, following curation of the TCGA dataset according to Peran *et al.* [26] to remove mischaracterized tumors from the cohort, overall survival was no longer significantly different, but disease-free survival was significantly improved (Fig 4b). These data suggest that the geneset was identifying patients with neuroendocrine tumors and the improved prognosis of these patients was influencing the overall survival results in the uncured dataset. To understand whether the geneset was associated with increased T cell infiltration, we obtained pre-calculated xCELL analysis of infiltrating cells in these pancreatic cancer TCGA specimens (<https://xcell.ucsf.edu>), and examined the correlation of each cell type with the expression of genes in our panel. To find potential patterns of biological significance, we correlated the expression of the gene set with the infiltration of immune cells across the TCGA cohort and performed clustering to gather co-regulated genes and cells together. We discovered that as with our cohort, there was a strong correlation between infiltrating T cells and dendritic cells in pancreatic tumors, and these correlated closely with the expression of genes in our panel (Fig 4c). Macrophage and endothelial cell infiltration did not correlate well with any of the genes in our panel, and smooth muscle cells, keratinocytes, and epithelial cell infiltration correlated best with some of the genes that were negatively associated with T cell infiltration, such as NEBL and HSP90AB1. These data indicate that the gene signature associated with high T cell infiltration in our pancreatic cancer cohort can similarly identify high T cell infiltration in other pancreatic cancer cohorts represented in the TCGA database. Further studies are needed to understand whether these genes have functional roles and are potential therapeutic targets.

Discussion

Despite the poor overall prognosis of pancreatic cancer, patients with high numbers of infiltrating T cells as determined by quantitative immunohistology have improved outcome. To understand the complexity of the tumor immune environment, we performed RNA-Seq and evaluated gene expression-based analyses of tumor-infiltrating cells. We found that there were limitations in current gene expression analyses of infiltrating immune cells, particularly where overall infiltration was low. Both CIBERSORT and xCELL showed poor concordance with IHC, and individual immune cell types identified by gene expression analysis had limited prognostic value. This could be somewhat overcome by aggregating molecularly-identified immune populations, for example into total T cell infiltrates, which showed some correlation to quantitative immunohistology and could be predictive of outcome. To determine if we could identify alternative molecular signatures of highly infiltrated tumors, we performed class comparison of gene expression and identified a transcriptional pattern in pancreatic adenocarcinoma that had predictive value for disease-free survival in the TCGA cohort.

Immunohistology with validated antibodies is the gold standard for quantitative assessment of infiltrating immune cells in cancer [27]. However, the diversity of cell types and limited number of cell-type specific markers makes it difficult to accurately assess many infiltrating cell types using histology. Flow cytometry is more able to address this diversity, using a series of gates to distinguish cell subtypes expressing multiple overlapping markers; however, this cannot be performed with archived tissues. The recent improvement in image analysis and technological improvements in multiplex staining has permitted much more complex assessment of tumors using immunohistology [28]. This is critical since many of our single markers have limitations. For example, our study used CD68 as a well-validated marker for tumor associated macrophages. However, there are significant limitations in the use of CD68 as a sole

marker of macrophages in tumors [20], particularly in view of the diverging phenotypes macrophages can generate. In particular, while there may be a spectrum of macrophage phenotypes [29], the polarized M1 (classical) versus M2 (alternative) macrophage phenotype has proven useful in discriminating macrophages that support versus suppress adaptive immunity to tumors [30–33]. Therefore, the presence of macrophages does not necessarily indicate that they generate immune suppression in the tumor.

The recent development of algorithms that can analyze gene expression data to estimate the prevalence of the broad range of cell types in a mixed tissue samples, combined with the increasing affordability of comprehensive genomic profiling of patient tumors, has opened new avenues of research [21]. While gene expression data can provide a great deal of information from small quantities of tissue, there are potential issues in estimating immune cell numbers in pancreatic tumors, since the abundance of some of these populations can be very low even when they have prognostic significance. For example, xCELL gene signatures were identified using purified populations and validated on peripheral blood samples [5], which have a very different immune cell abundance when compared to tumors. CIBERSORT was shown to have superior performance to other approaches available at the time to assess immune infiltrating cells from genomic data [6]; however, performance was limited when the target immune populations represented fewer than 1% of the spiked mixture. Each approach provides valuable information on immune infiltration and may be sufficient to assess the environment of more abundantly infiltrated tumors. However, there are limitations based on the number of RNA reads provided by each infiltrating cell in a bulk population. Immunohistology has its own limitations, particularly those relating to epitope preservation through tissue processing, and the difficulties in standardizing staining over time and between institutions. In this study we set quantitative immunohistology as the gold standard for comparison; however, standard sampling issues such as selection of an appropriate archived tissue block and the relevance of a single 5 μ M section to the tissue as a whole can lead to inaccuracies that apply to each approach. Novel technologies are emerging that incorporate the geographic information of histology with comprehensive gene profiling, and have the potential to change how we assess the immune complexity of tumors [34]. Further analysis of patients giving discordant RNA and IHC data would be valuable to understand the impact of sampling versus other complicating factors that could explain the variations.

The strength of genomic and other omic profiling is the wealth of data that can be extracted simultaneously. Along with infiltrating cells, omic analyses can inform as to the mutational status of the tumor [8] to identify immunotherapeutic targets [35], and identify expression of inflammatory and chemokine markers that may dictate immune cell recruitment [36, 37]. Such analyses would be best combined with genomic analysis approaches that subdivide patients according to novel molecular subtypes, some of which include tumor subtypes associated with higher immune infiltrates [38, 39]. However, further subdivision of patients will require much larger cohorts to generate meaningful results. All of this can be performed on very small quantities of patient material, at decreasing cost and with increasing speed. While IHC-based multiplexing continues to increase the number of analytes that can be assessed on a single tissue sample, this approach depends on the availability of high-quality validated antibodies for each target. Unbiased sequencing-based approaches are to some degree future-proofed against genes that may be of interest yet do not currently have reagents for IHC or other analyses.

Further refinement of gene expression profiling for pancreatic adenocarcinoma and similar poorly infiltrated tumors could have significant benefits in personalizing immunotherapy for these recalcitrant tumors. For example, we show that CTLA4 is enriched in patients with high

T cell infiltration and is part of the gene set that is associated with improved disease free survival. Antibodies targeting CTLA4 are an effective immunotherapy in some tumors [40], but single agent anti-CTLA4 is not effective in patients with locally advanced and metastatic pancreatic adenocarcinoma [41]. In preclinical models of pancreatic adenocarcinoma, we similarly found that anti-CTLA4 is ineffective as a single agent [42]. The combination of anti-CTLA4 and radiation therapy is curative, but only where the host has good pre-existing immunity to the tumor. Thus, gene expression profiling may help identify patient subsets with adequate T cell infiltration that may benefit from immunotherapy combinations, and direct other patients to novel interventions to improve their tumor immune environment prior to further treatment.

Conclusions

While immunotherapy options are currently limited for patients with pancreatic adenocarcinoma, these and other data showing an impact of immune infiltrates on patient outcomes suggests we should continue to refine our understanding of the immune environment and pursue immune therapies that are appropriate to the particular tumor environments of pancreatic cancer. At present RNASeq-based analyses must take into account the poor overall infiltrate in some tumor types to provide accurate assessments of the tumor environment.

Supporting information

S1 Fig. Pearson correlation matrix of infiltrating immune cells calculated from RNASeq data using CIBERSORT. Rows are centered; no scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage.
(PDF)

S2 Fig. Pearson correlation matrix of infiltrating immune cells calculated from RNASeq data using xCELL. Rows are centered; no scaling is applied to rows. Both rows and columns are clustered using Manhattan distance and average linkage.
(PDF)

S3 Fig. Heatmap showing percent match between genes used in xCell gene signature for each defined cell type. Rows are centered; no scaling is applied to rows. Both rows and columns are clustered using correlation distance and average linkage.
(PDF)

S4 Fig. Comparison of all RNA-based approaches to assess infiltrates. a) The immune infiltrate score determined by MCPcounter of i) MCPcounter T cell vs. IHC CD3⁺, ii) MCPcounter CD8 T cell vs IHC CD8⁺, and iii) MCPcounter monocytic lineage vs IHC CD68⁺ cell populations from the same patient. Each symbol represents one patient. b) analysis as in a) for i) EPIC Bref CD8 T cells and ii) EPIC Tref macrophages. c) Pearson correlation matrix of infiltrating T cell subsets calculated from RNASeq data using xCELL, CIBERSORT, MCPcounter, EPIC Bref, and EPIC Tref. Rows are centered; no scaling is applied to rows. Both rows and columns are clustered using Manhattan distance and average linkage.
(PDF)

S1 Table. Gene-based expression levels for all patients with RNASeq analysis in the study, along with quantitative IHC.
(TXT)

Author Contributions

Conceptualization: Brady Bernard, Brian Piening, Pippa Newell, Paul Hansen, Kristina Young, Carlo Bifulco, Marka R. Crittenden, Michael J. Gough.

Data curation: Michael J. Gough.

Funding acquisition: Zeljka Jutric, Ephraim Tang, Marka R. Crittenden, Michael J. Gough.

Investigation: Brady Bernard, Venkatesh Rajamanickam, Christopher Dubay, Brian Piening, Emilio Alonso, Zeljka Jutric, Ephraim Tang, Terry Medler, Andrew Gunderson, Joanna Pucliowska, Michael J. Gough.

Methodology: Brady Bernard, Venkatesh Rajamanickam, Christopher Dubay, Brian Piening, Terry Medler, Andrew Gunderson, Carlo Bifulco, Joanna Pucliowska, Marka R. Crittenden, Michael J. Gough.

Project administration: Michael J. Gough.

Resources: Brady Bernard, Brian Piening, Emilio Alonso, Pippa Newell, Paul Hansen, Kristina Young, Carlo Bifulco, Marka R. Crittenden, Michael J. Gough.

Supervision: Brady Bernard, Michael J. Gough.

Visualization: Brady Bernard, Venkatesh Rajamanickam, Christopher Dubay, Michael J. Gough.

Writing – original draft: Michael J. Gough.

Writing – review & editing: Brady Bernard, Brian Piening, Marka R. Crittenden.

References

1. Tang ES, Newell PH, Wolf RF, Hansen PD, Cottam B, Ballesteros-Merino C, et al. Association of Immunologic Markers With Survival in Upfront Resectable Pancreatic Cancer. *JAMA surgery*. 2018; 153(11):1055–7. <https://doi.org/10.1001/jamasurg.2018.1757> PMID: 30090929
2. Miksch RC, Schoenberg MB, Weniger M, Bosch F, Ormanns S, Mayer B, et al. Prognostic Impact of Tumor-Infiltrating Lymphocytes and Neutrophils on Survival of Patients with Upfront Resection of Pancreatic Cancer. *Cancers (Basel)*. 2019; 11(1).
3. Nejati R, Goldstein JB, Halperin DM, Wang H, Hejazi N, Rashid A, et al. Prognostic Significance of Tumor-Infiltrating Lymphocytes in Patients With Pancreatic Ductal Adenocarcinoma Treated With Neoadjuvant Chemotherapy. *Pancreas*. 2017; 46(9):1180–7. <https://doi.org/10.1097/MPA.0000000000000914> PMID: 28902789
4. Fukunaga A, Miyamoto M, Cho Y, Murakami S, Kawarada Y, Oshikiri T, et al. CD8+ tumor-infiltrating lymphocytes together with CD4+ tumor-infiltrating lymphocytes and dendritic cells improve the prognosis of patients with pancreatic adenocarcinoma. *Pancreas*. 2004; 28(1):e26–31. <https://doi.org/10.1097/00006676-200401000-00023> PMID: 14707745
5. Aran D, Hu Z, Butte AJ. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol*. 2017; 18(1):220. <https://doi.org/10.1186/s13059-017-1349-1> PMID: 29141660
6. Newman AM, Liu CL, Green MR, Gentles AJ, Feng W, Xu Y, et al. Robust enumeration of cell subsets from tissue expression profiles. *Nat Methods*. 2015; 12(5):453–7. <https://doi.org/10.1038/nmeth.3337> PMID: 25822800
7. Thorsson V, Gibbs DL, Brown SD, Wolf D, Bortone DS, Ou Yang TH, et al. The Immune Landscape of Cancer. *Immunity*. 2018; 48(4):812–30 e14. <https://doi.org/10.1016/j.immuni.2018.03.023> PMID: 29628290
8. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med*. 2017; 9(1):34. <https://doi.org/10.1186/s13073-017-0424-2> PMID: 28420421
9. Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods*. 2017; 14(4):417–9. <https://doi.org/10.1038/nmeth.4197> PMID: 28263959

10. Metsalu T, Vilo J. ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res.* 2015; 43(W1):W566–70. <https://doi.org/10.1093/nar/gkv468> PMID: 25969447
11. Kimes PK, Liu Y, Neil Hayes D, Marron JS. Statistical significance for hierarchical clustering. *Bio-metrics.* 2017; 73(3):811–21. <https://doi.org/10.1111/biom.12647> PMID: 28099990
12. Wright GW, Simon RM. A random variance model for detection of differential gene expression in small microarray experiments. *Bioinformatics.* 2003; 19(18):2448–55. <https://doi.org/10.1093/bioinformatics/btg345> PMID: 14668230
13. Korn EL, Troendle JF, McShane LM, Simon R. Controlling the number of false discoveries: application to high-dimensional genomic data. *Journal of Statistical Planning and Inference.* 2004; 124(2):379–98.
14. Wu G, Dawson E, Duong A, Haw R, Stein L. ReactomeFIViz: a Cytoscape app for pathway and network-based data analysis. *F1000Res.* 2014; 3:146. <https://doi.org/10.12688/f1000research.4431.2> PMID: 25309732
15. Liu J, Lichtenberg T, Hoadley KA, Poisson LM, Lazar AJ, Cherniack AD, et al. An Integrated TCGA Pan-Cancer Clinical Data Resource to Drive High-Quality Survival Outcome Analytics. *Cell.* 2018; 173(2):400–16 e11. <https://doi.org/10.1016/j.cell.2018.02.052> PMID: 29625055
16. Cerami E, Gao J, Dogrusoz U, Gross BE, Sumer SO, Aksoy BA, et al. The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data. *Cancer discovery.* 2012; 2(5):401–4. <https://doi.org/10.1158/2159-8290.CD-12-0095> PMID: 22588877
17. Gao J, Aksoy BA, Dogrusoz U, Dresdner G, Gross B, Sumer SO, et al. Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal. *Sci Signal.* 2013; 6(269):pl1. <https://doi.org/10.1126/scisignal.2004088> PMID: 23550210
18. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014; 15(12):550. <https://doi.org/10.1186/s13059-014-0550-8> PMID: 25516281
19. Mitchem JB, Brennan DJ, Knolhoff BL, Belt BA, Zhu Y, Sanford DE, et al. Targeting tumor-infiltrating macrophages decreases tumor-initiating cells, relieves immunosuppression, and improves chemotherapeutic responses. *Cancer Res.* 2013; 73(3):1128–41. <https://doi.org/10.1158/0008-5472.CAN-12-2731> PMID: 23221383
20. Ruffell B, Au A, Rugo HS, Esserman LJ, Hwang ES, Coussens LM. Leukocyte composition of human breast cancer. *Proc Natl Acad Sci U S A.* 2012; 109(8):2796–801. <https://doi.org/10.1073/pnas.1104303108> PMID: 21825174
21. Gentles AJ, Newman AM, Liu CL, Bratman SV, Feng W, Kim D, et al. The prognostic landscape of genes and infiltrating immune cells across human cancers. *Nat Med.* 2015; 21(8):938–45. <https://doi.org/10.1038/nm.3909> PMID: 26193342
22. Racle J, de Jonge K, Baumgaertner P, Speiser DE, Gfeller D. Simultaneous enumeration of cancer and immune cell types from bulk tumor gene expression data. *Elife.* 2017;6.
23. Becht E, Giraldo NA, Lacroix L, Buttard B, Elarouci N, Petitprez F, et al. Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression. *Genome Biol.* 2016; 17(1):218. <https://doi.org/10.1186/s13059-016-1070-5> PMID: 27765066
24. Peggs KS, Quezada SA, Korman AJ, Allison JP. Principles and use of anti-CTLA4 antibody in human cancer immunotherapy. *Curr Opin Immunol.* 2006; 18(2):206–13. <https://doi.org/10.1016/j.coi.2006.01.011> PMID: 16464564
25. Miller BC, Sen DR, Al Abosy R, Bi K, Virkud YV, LaFleur MW, et al. Subsets of exhausted CD8(+) T cells differentially mediate tumor control and respond to checkpoint blockade. *Nat Immunol.* 2019; 20(3):326–36. <https://doi.org/10.1038/s41590-019-0312-6> PMID: 30778252
26. Peran I, Madhavan S, Byers SW, McCoy MD. Curation of the Pancreatic Ductal Adenocarcinoma Subset of the Cancer Genome Atlas Is Essential for Accurate Conclusions about Survival-Related Molecular Mechanisms. *Clin Cancer Res.* 2018; 24(16):3813–9. <https://doi.org/10.1158/1078-0432.CCR-18-0290> PMID: 29739787
27. Pages F, Mlecnik B, Marliot F, Bindea G, Ou FS, Bifulco C, et al. International validation of the consensus Immunoscore for the classification of colon cancer: a prognostic and accuracy study. *Lancet.* 2018; 391(10135):2128–39. [https://doi.org/10.1016/S0140-6736\(18\)30789-X](https://doi.org/10.1016/S0140-6736(18)30789-X) PMID: 29754777
28. Wickenhauser C, Bethmann D, Feng Z, Jensen SM, Ballesteros-Merino C, Massa C, et al. Multispectral Fluorescence Imaging Allows for Distinctive Topographic Assessment and Subclassification of Tumor-Infiltrating and Surrounding Immune Cells. *Methods Mol Biol.* 2019; 1913:13–31. https://doi.org/10.1007/978-1-4939-8979-9_2 PMID: 30666596
29. Xue J, Schmidt SV, Sander J, Draffehn A, Krebs W, Quester I, et al. Transcriptome-based network analysis reveals a spectrum model of human macrophage activation. *Immunity.* 2014; 40(2):274–88. <https://doi.org/10.1016/j.immuni.2014.01.006> PMID: 24530056

30. Rodriguez PC, Quiceno DG, Zabaleta J, Ortiz B, Zea AH, Piazuelo MB, et al. Arginase I production in the tumor microenvironment by mature myeloid cells inhibits T-cell receptor expression and antigen-specific T-cell responses. *Cancer Res.* 2004; 64(16):5839–49. <https://doi.org/10.1158/0008-5472.CAN-04-0465> PMID: 15313928
31. Bronte V, Serafini P, De Santo C, Marigo I, Tosello V, Mazzone A, et al. IL-4-induced arginase 1 suppresses alloreactive T cells in tumor-bearing mice. *J Immunol.* 2003; 170(1):270–8. <https://doi.org/10.4049/jimmunol.170.1.270> PMID: 12496409
32. Mantovani A, Sica A, Sozzani S, Allavena P, Vecchi A, Locati M. The chemokine system in diverse forms of macrophage activation and polarization. *Trends Immunol.* 2004; 25(12):677–86. <https://doi.org/10.1016/j.it.2004.09.015> PMID: 15530839
33. Mantovani A, Sozzani S, Locati M, Allavena P, Sica A. Macrophage polarization: tumor-associated macrophages as a paradigm for polarized M2 mononuclear phagocytes. *Trends Immunol.* 2002; 23:549–55. [https://doi.org/10.1016/s1471-4906\(02\)02302-5](https://doi.org/10.1016/s1471-4906(02)02302-5) PMID: 12401408
34. Berglund E, Maaskola J, Schultz N, Friedrich S, Marklund M, Bergenstrahle J, et al. Spatial maps of prostate cancer transcriptomes reveal an unexplored landscape of heterogeneity. *Nature communications.* 2018; 9(1):2419. <https://doi.org/10.1038/s41467-018-04724-5> PMID: 29925878
35. Tran E, Ahmadzadeh M, Lu YC, Gros A, Turcotte S, Robbins PF, et al. Immunogenicity of somatic mutations in human gastrointestinal cancers. *Science.* 2015; 350(6266):1387–90. <https://doi.org/10.1126/science.aad1253> PMID: 26516200
36. Harlin H, Meng Y, Peterson AC, Zha Y, Tretiakova M, Slingluff C, et al. Chemokine expression in melanoma metastases associated with CD8+ T-cell recruitment. *Cancer Res.* 2009; 69(7):3077–85. <https://doi.org/10.1158/0008-5472.CAN-08-2281> PMID: 19293190
37. Li J, Byrne KT, Yan F, Yamazoe T, Chen Z, Baslan T, et al. Tumor Cell-Intrinsic Factors Underlie Heterogeneity of Immune Cell Infiltration and Response to Immunotherapy. *Immunity.* 2018; 49(1):178–93 e7. <https://doi.org/10.1016/j.immuni.2018.06.006> PMID: 29958801
38. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature.* 2016; 531(7592):47–52. <https://doi.org/10.1038/nature16965> PMID: 26909576
39. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat Med.* 2011; 17(4):500–3. <https://doi.org/10.1038/nm.2344> PMID: 21460848
40. Hodi FS, O'Day SJ, McDermott DF, Weber RW, Sosman JA, Haanen JB, et al. Improved survival with ipilimumab in patients with metastatic melanoma. *N Engl J Med.* 2010; 363(8):711–23. <https://doi.org/10.1056/NEJMoa1003466> PMID: 20525992
41. Royal RE, Levy C, Turner K, Mathur A, Hughes M, Kammula US, et al. Phase 2 trial of single agent Ipilimumab (anti-CTLA-4) for locally advanced or metastatic pancreatic adenocarcinoma. *J Immunother.* 2010; 33(8):828–33. <https://doi.org/10.1097/CJI.0b013e3181eec14c> PMID: 20842054
42. Crittenden MR, Zebertavage L, Kramer G, Bambina S, Friedman D, Troesch V, et al. Tumor cure by radiation therapy and checkpoint inhibitors depends on pre-existing immunity. *Sci Rep.* 2018; 8(1):7012. <https://doi.org/10.1038/s41598-018-25482-w> PMID: 29725089