



OPEN ACCESS

EDITED BY

Wei Wang,
Edith Cowan University, Australia

REVIEWED BY

Ekene Emmanuel Nweke,
University of the Witwatersrand, South Africa
Geeta Upadhyay,
Uniformed Services University of the Health Sciences, United States

*CORRESPONDENCE

Oliver F. Bathe,
✉ oliver.bathe@albertahealthservices.ca

RECEIVED 24 August 2023

ACCEPTED 06 October 2023

PUBLISHED 30 October 2023

CITATION

Pervin J, Asad M, Cao S, Jang GH, Feizi N, Haibe-Kains B, Karasinska JM, O'Kane GM, Gallinger S, Schaeffer DF, Renouf DJ, Zogopoulos G and Bathe OF (2023), Clinically impactful metabolic subtypes of pancreatic ductal adenocarcinoma (PDAC). *Front. Genet.* 14:1282824. doi: 10.3389/fgene.2023.1282824

COPYRIGHT

© 2023 Pervin, Asad, Cao, Jang, Feizi, Haibe-Kains, Karasinska, O'Kane, Gallinger, Schaeffer, Renouf, Zogopoulos and Bathe. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Clinically impactful metabolic subtypes of pancreatic ductal adenocarcinoma (PDAC)

Jannat Pervin¹, Mohammad Asad², Shaolong Cao³, Gun Ho Jang⁴, Nikta Feizi⁵, Benjamin Haibe-Kains⁵, Joanna M. Karasinska⁶, Grainne M. O'Kane⁷, Steven Gallinger⁴, David F. Schaeffer⁸, Daniel J. Renouf⁹, George Zogopoulos¹⁰ and Oliver F. Bathe^{1,11*}

¹Department of Oncology, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada, ²Department of Biochemistry and Molecular Biology, University of Calgary, Calgary, AB, Canada, ³Department of Bioinformatics and Computational Biology, University of Texas MD Anderson Cancer Centre, Houston, TX, United States, ⁴Ontario Institute for Cancer Research, Toronto, ON, Canada, ⁵Princess Margaret Cancer Centre, University Health Network, Toronto, ON, Canada, ⁶Pancreas Centre BC, Vancouver, BC, Canada, ⁷University Health Network, University of Toronto, Toronto, ON, Canada, ⁸Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC, Canada, ⁹Department of Medicine, University of British Columbia, Vancouver, BC, Canada, ¹⁰Department of Surgery, McGill University Health Centre, McGill University, Montreal, QC, Canada, ¹¹Department of Surgery, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada

Background: Pancreatic ductal adenocarcinoma (PDAC) is a lethal disease characterized by a diverse tumor microenvironment. The heterogeneous cellular composition of PDAC makes it challenging to study molecular features of tumor cells using extracts from bulk tumor. The metabolic features in tumor cells from clinical samples are poorly understood, and their impact on clinical outcomes are unknown. Our objective was to identify the metabolic features in the tumor compartment that are most clinically impactful.

Methods: A computational deconvolution approach using the DeMixT algorithm was applied to bulk RNASeq data from The Cancer Genome Atlas to determine the proportion of each gene's expression that was attributable to the tumor compartment. A machine learning algorithm designed to identify features most closely associated with survival outcomes was used to identify the most clinically impactful metabolic genes.

Results: Two metabolic subtypes (M1 and M2) were identified, based on the pattern of expression of the 26 most important metabolic genes. The M2 phenotype had a significantly worse survival, which was replicated in three external PDAC cohorts. This PDAC subtype was characterized by net glycogen catabolism, accelerated glycolysis, and increased proliferation and cellular migration. Single cell data demonstrated substantial intercellular heterogeneity in the metabolic features that typified this aggressive phenotype.

Conclusion: By focusing on features within the tumor compartment, two novel and clinically impactful metabolic subtypes of PDAC were identified. Our study emphasizes the challenges of defining tumor phenotypes in the face of the significant intratumoral heterogeneity that typifies PDAC. Further studies are

required to understand the microenvironmental factors that drive the appearance of the metabolic features characteristic of the aggressive M2 PDAC phenotype.

KEYWORDS

pancreatic ductal adenocarcinoma, pancreatic cancer, metabolism, deconvolution, prognosis

Introduction

Pancreatic adenocarcinoma (PDAC) is highly lethal, with a 5-year survival <5% in unresectable cases. Even after curative intent resection, tumor recurs in the majority of instances. While there have been advances in systemic therapies for PDAC, only a small impact on clinical outcomes has been realized. To better understand PDAC biology, several groups have described molecular subgroups with unique clinical and biological features (Collisson et al., 2011; Moffitt et al., 2015; Bailey et al., 2016; Raphael and Cancer Genome Atlas Research Network, 2017; Cao et al., 2021). Most of these efforts have employed bulk (whole tumor) analysis (Collisson et al., 2011; Bailey et al., 2016; Raphael and Cancer Genome Atlas Research Network, 2017; Cao et al., 2021). However, PDAC characteristically has a prominent stromal component consisting of diverse cellular populations including fibroblasts, stellate cells, acinar cells, endothelial cells, and immune cells. The situation is further complicated by highly variable stromal features. While this can be technically addressed by single cell analyses, so far series describing single cell analyses have been too small to provide insight on the effects of individual cellular components on clinical outcomes (Wang et al., 2021; Fang et al., 2022).

It is well known that PDAC afflicts the host with significant metabolic perturbations manifested as impaired glycemic control (Alpertunga et al., 2021; Jeon et al., 2021) and cachexia (Rupert et al., 2021; Zhong et al., 2022). Studies on the circulating metabolome demonstrate distinct features in comparison to controls (Bathe et al., 2011; Mayers et al., 2014), although so far metabolomic subtypes based on the circulating metabolome have not been described. Three metabolic subtypes based on metabolomic profiles have been described in cell lines, including glycolytic and lipogenic subtypes and a subtype with reduced proliferative capacity (Daemen et al., 2015). The metabolic subtypes had differential sensitivities to inhibitors of glycolysis and glutaminolysis. Based on these findings, Karasinska et al. (2020) evaluated the association of glycolytic and cholesterogenic gene expression levels on survival. This targeted analysis demonstrated that glycolytic tumors had the shortest median survival, and cholesterogenic PDACs had the longest survival.

While each of these studies has shed light on very important metabolic features, it is unclear what metabolic features specifically in tumor cells are most impactful on the biology and clinical behavior of PDAC. We postulated that subsets of PDAC had metabolic features that were clinically impactful and potentially actionable from a therapeutic perspective. To separate the effects of the tumor cell compartment from the stromal compartment, we employed a computational deconvolution approach. The most clinically impactful metabolic features of PDAC were identified

using a proprietary machine learning algorithm (HighLifeR™) (Craig et al., 2022) designed to expose features that were most closely associated with survival outcomes.

Materials and methods

Patient samples and data acquisition

In this study, PDAC cases annotated by The Cancer Genome Atlas (TCGA) ($N = 142$) were used as discovery cohort. The normalized RSEM RNASeq data were downloaded from firebrowse.org, and related clinical information was obtained from the Genomic Data Commons Data portal (GDC PDAC). To derive gene expression levels in the tumor compartment, bulk RNASeq data from TCGA were submitted to computational deconvolution. A two-compartment model was employed using the R package DeMixT (Wang et al., 2018; Cao et al., 2022). The discovery cohort consisted of 12,635 protein-coding genes, including 1,499 metabolic genes. The metabolic gene list was aggregated from the Reactome Pathway Database Vol 77. Genes belonging to “Metabolism; Id: R-HAS-143072B.10, Species: *Homo sapiens*.” Data on mutations and copy number variations (hg38) were downloaded using the TCGAbiolinks package in R. Previously defined hypoxia scores of PDAC bulk tumors were downloaded from cBioPortal.

To validate the findings from the discovery cohort, HTSeq RNASeq data and clinical information were obtained from CPTAC (using the R package TCGAbiolinks from the GDC portal), ICGC (from the ICGC Data Portal), and from the COMPASS trial (Aung et al., 2018; Zhang et al., 2019; Cao et al., 2021). Each dataset was independently normalized using the median of ratios, and z -scores of all genes were calculated using the scale function in R. A prediction model based on the prognostic metabolic genes was generated using the WEKA software (Witten et al., 2017). The multilayer perceptron (MLP) algorithm, a form of artificial neural network (Murtagh, 1991), demonstrated the highest accuracy on internal validation.

Identification of prognostic genes using HighLifeR™

HighLifeR™ (Qualisure Diagnostics Inc., Calgary, Canada) facilitates the interrogation of highly dimensional datasets with relatively limited sample size, to identify features that are most closely associated with survival events. HighLifeR™ employs partial Cox regression using predictions from the latent components method introduced by Li and Gui (2004). The algorithm involves the recursive application of Cox proportional hazards estimates, testing a multitude of genes and patient combinations in a

supervised machine learning context. It also employs randomization of training samples into “virtual cohorts” for 20 testing cycles, each including roughly 70% of patients, with resampling to minimize outlier effects. The HighLifeR™ statistical mechanism includes: a) extensive combinatorial iterations to establish gene prognosis in variable space; b) selection of the most highly ranked genes with respect to their association with survival; c) development of a combined prognostic rating model. Prognostic genes are chosen based on their recurring top-200 ranking and prognostic influence (Wald statistic) exceeding half of the maximum Wald statistic in the training set (>4.7). Broad utilization of randomization (including in sample distribution for testing and validation, and in combination sequences) reduces the chance of identifying a sample set-bound prognostic pattern.

Identification of metabolic subgroups

Selected prognostic metabolic genes ($N = 26$) identified using HighLifeR™ were submitted to unsupervised analysis to uncover underlying patterns. The k-means clustering algorithm was utilized using the Euclidean distance metric in conjunction with the complete linkage clustering method, with the number of clusters (k) ranging from 2 to 6. Finally, $k = 2$ clusters were selected to define the metabolic subtypes, denoted as M1 and M2. For visualization, Euclidean distance metric with complete linkage clustering was employed using the ComplexHeatmap package in R. To conduct the survival analysis, Kaplan-Meier plots were generated using GraphPad PRISM version 8.0.

Mutation and copy number variation analysis

Mutational and copy number variation analysis was carried out using the R package “maftools.” For somatic copy number variations, GISTIC2.0 was used to download the relevant data. A Fisher’s exact test was used to compare mutation frequencies and copy number variations between groups. To account for multiple comparisons, p -values were adjusted using the Benjamin-Hochberg method.

Differentially expressed genes analysis

Differentially expressed genes (DEGs) were identified using the limma-voom function in R (Law et al., 2014). Briefly, data were transformed to log₂ counts per million reads (CPM). Genes with less than 1 CPM were excluded from the analysis of differentially expressed genes. The metabolic subtype with better survival (M1) was used as the reference group. We adjusted for multiple comparisons using the Benjamin-Hochberg method. The significance threshold for this analysis was an adjusted p -value of ≤ 0.05 and log fold changes of $\geq \pm 1$. Ultimately, 551 DEGs were identified based on these criteria.

Gene set enrichment and pathway analysis

Gene set enrichment analysis (GSEA) was performed using version 4.1.0 of the Broad Institute GSEA software. Using all

protein coding genes, the 50 Hallmarks version 7.5 gene sets from the Broad Institute were interrogated. Then a focused evaluation for enrichment of metabolism-related pathways was performed, limiting the input to the 1,499 metabolic genes present in the expression dataset from the discovery cohort. The 46 pathways dedicated to metabolism were sourced from the Reactome database (data available at <https://reactome.org/PathwayBrowser/#/R-HSA-1430728>). We determined significance levels through a permutation-based approach and adjusted for multiple testing using the Benjamini-Hochberg method to calculate the false discovery rate (FDR). The enrichment score was computed using a ranking-based metric, which measures the cumulative distribution of genes within a gene set relative to the entire dataset. Results with a significance cutoff of $p < 0.05$ and an FDR threshold of <25% were considered statistically significant.

For the IPA (Ingenuity Pathway Analysis) analysis, a total of 551 DEGs, along with their adjusted p -values and fold changes, were used as input. The IPA analysis was conducted using the Qiagen IPA platform (<https://digitalinsights.qiagen.com/products-overview/discovery-insights-portfolio/analysis-and-visualization/qiagen-ipa/>). Fisher’s exact test was utilized to assess the statistical significance of pathway and function enrichment. We controlled for multiple testing using the Benjamin-Hochberg correction and calculated activated z -scores to predict the likely activation state of biological processes based on DEGs. A significance cut-off of nominal p -value 0.05 was applied for the canonical pathways, while a Benjamin-Hochberg (BH) p -value of 0.05 was used for the functional analysis. Pathway visualizations were generated using various R packages.

Single cell analysis

To explore single-cell characteristics of PDAC primary tumors and metastasis samples, we retrieved the NCBI-GEO dataset with the ID GSE154778 (Lin, W. et al., 2020). The dataset included single cell RNAseq data from 10 primary PDACs and 6 samples from metastases. The RNA-seq data for each sample were obtained following the methodology described in the study. The focus of the downstream analysis was on epithelial tumor cells (ETCs) to validate the presence of inter-tumor heterogeneity. Classification of ETCs was based on marker genes, including EpCAM and KRT19. Log-normalized data and z -scores for prognostic metabolic genes were calculated for each tumor sample.

Drug sensitivity testing

Pharmacogenomic profiles from drug screening studies of PDAC cell lines from CCLE and gCSI were downloaded from Orcestra.ca (<https://orcestra.ca/pset>). Using the MLP model, cell lines were classified as M1 or M2. The data from CCLE and gCSI were not combined to mitigate potential confounding factors such as variations in drug concentration range or cell viability assay methods. To compare the drug sensitivity status between the two subtypes, the Wilcoxon signed-rank test was used to

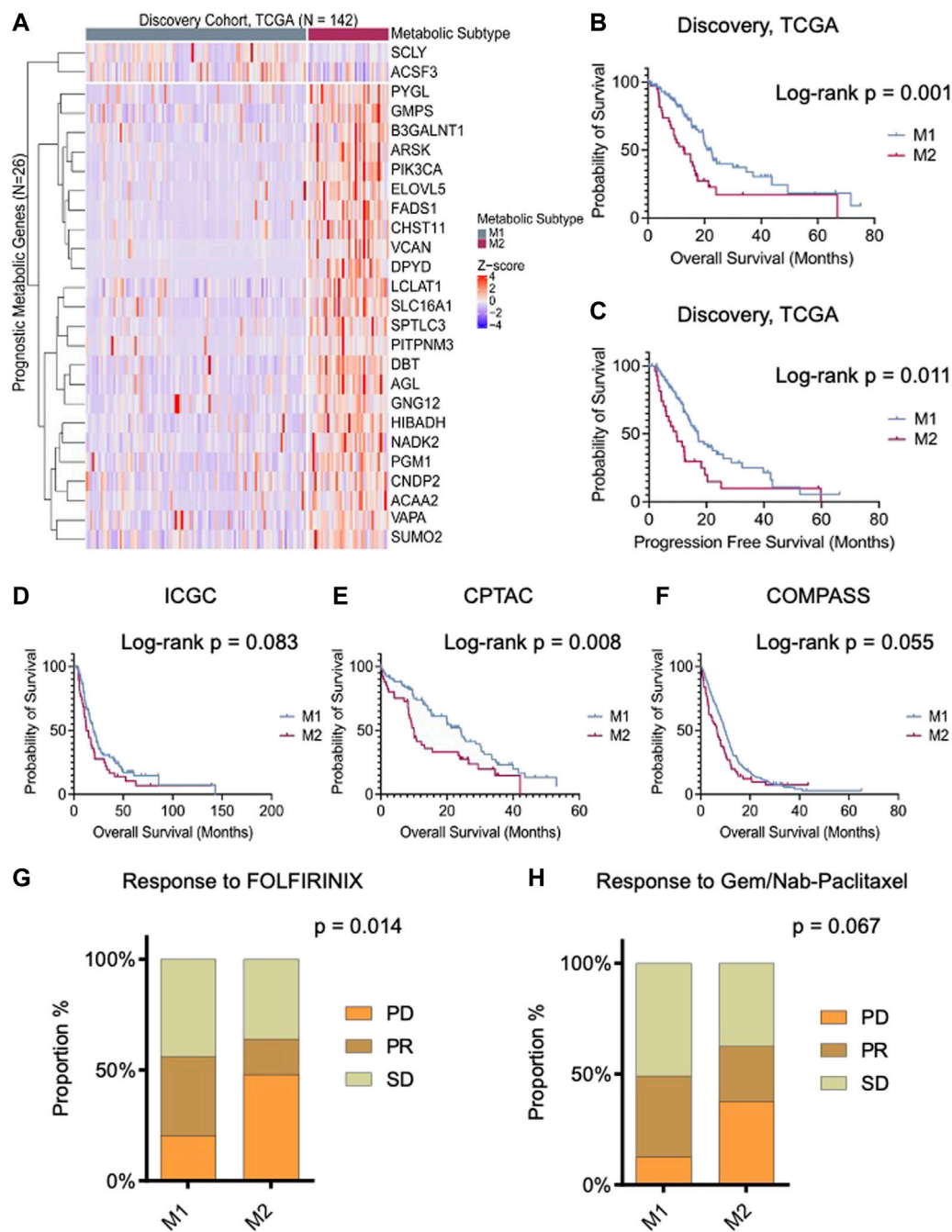


FIGURE 1

Prognostic tumor metabolic subtypes in pancreatic ductal adenocarcinoma (PDAC). (A) Unsupervised clustering ($k = 2$) illustrating distinct metabolic subgroups in the TCGA PDAC cohort ($N = 142$). The heatmap depicts expression levels of prognostic metabolic genes within the identified subtypes, highlighting both “lower-risk” (M1) and “high-risk” (M2) groups. (B,C) Kaplan-Meier plot displaying the overall and progression-free survival of the metabolic subtypes in the discovery cohort, demonstrating significant differences in survival outcomes. (D–F) Overall survival analysis of resectable PDAC patients: International Cancer Genome Consortium (ICGC, $n = 172$), Clinical Proteomic Tumor Analysis Consortium (CPTAC, $n = 140$), and unresectable patients from the COMPASS cohort ($n = 272$) stratified according to predicted metabolic subtypes. (G,H) Bar chart presents the responses to the first line treatment in COMPASS cohort. M1 subtype exhibits stable and partial responses to both FOLFIRINOX and Gemcitabine/Nab-paclitaxel. On the contrary, M2 group shows poor response to gemcitabine and combined therapy. Log-rank and Chi-square p -values are shown in the figure.

compare the distribution of the area above the dose-response curve (AAC) values per drug across the subtypes. p -values were adjusted using the Benjamini-Hochberg method to account for multiple testing.

Additional statistical details

In addition to the analyses described above, continuous variables were compared using the Student’s t -test, and the Mann-Whitney U

TABLE 1 Patient characteristics in TCGA cases with each metabolic subtype phenotype.

Characteristics	PDAC tumor metabolic subtypes		<i>p</i> -value
	M1 (Lower-risk) No (<i>n</i> = 104) (%)	M2 (High-risk) No (<i>n</i> = 38) (%)	
Age (years) Mean ± SD	64.54 ± 10.633	66.58 ± 10.892	0.440
Sex			0.684
Female	48 (46.2)	19 (50.0)	
Male	56 (53.8)	19 (50.0)	
Race			0.761
Asian	5 (5.0)	3 (7.9)	
African American	4 (4.0)	1 (2.6)	
White	91 (91.0)	34 (89.5)	
Histologic Grade			0.219
G1	16 (15.4)	4 (10.5)	
G2	60 (57.7)	19 (50.0)	
G3	28 (26.9)	14 (36.8)	
G4	0 (0.0)	1 (2.6)	
Tumour Stage			0.717
T1	2 (1.9)	1 (2.7)	
T2	13 (12.5)	3 (8.1)	
T3	87 (83.7)	33 (89.2)	
T4	2 (1.9)	0 (0.0)	
Lymph Node Stage			0.459
N0	28 (27.2)	8 (21.1)	
N1	75 (72.8)	30 (78.9)	
Metastatic Stage			0.871
M0	47 (94.0)	19 (95.0)	
M1	3 (6.0)	1 (5.0)	
Smoking History Category			0.538
1	33 (38.8)	17 (51.5)	
2	12 (14.1)	5 (15.2)	
3	21 (24.7)	4 (12.1)	
4	14 (16.5)	6 (18.2)	
5	5 (5.9)	1 (3.0)	
Alcohol History			0.830
No	37 (38.1)	13 (36.1)	
Yes	60 (61.9)	23 (63.9)	
Diabetes History			0.270
No	61 (70.9)	29 (80.6)	
Yes	25 (29.1)	7 (19.4)	

test was used for non-parametric data. Proportions were compared using the chi-square test. Log-rank tests were used to compare survivals between groups. Statistical analyses were executed using RStudio, GraphPad PRISM version 8.0, and SPSS version 26.

Results

Identification of two prognostic metabolic phenotypes

We used data from The Cancer Genome Atlas (TCGA) PDAC cohort consisting of 142 patients as our discovery dataset. The

majority of the TCGA PDAC samples were resected, and RNASeq was performed in bulk samples from primary tumors. The DeMixT algorithm was applied to determine the expression level of each gene that was attributable to the tumor compartment. HighLifeR™, a proprietary machine-learning algorithm designed to identify features that most strongly contribute to survival outcomes, was applied to the deconvolved tumor-specific expression of 1,499 genes with a known metabolic function. Briefly, HighLifeR™ calculates the Cox proportional hazards in an exhaustive series of permutations consisting of combinations of genes and patients. A set of 26 genes was identified as most consistently and significantly associated with overall survival (OS). Genes had diverse metabolic functions, including roles in

carbohydrate, amino acid, and lipid metabolism (Supplementary Table S1). Interestingly, dihydropyrimidine dehydrogenase (DPYD), the rate-limiting enzyme responsible for the metabolism of 5-fluorouracil, emerged in the list of prognostic metabolic genes. Tumor expression of DPYD has been linked to resistance to fluoropyrimidines (Gustavsson et al., 2009; Zhang et al., 2017). Upon searching the Human Protein Atlas, six of these genes (SPTLC3, PYGL, GMPS, ACSF3, SLC16A1, and PGM1) were confirmed to be prognostic in PDAC, based on expression at the protein level.

An unsupervised analysis focused on the 26 prognostic genes revealed a pattern of co-expression. Two clusters were apparent (M1 and M2; Figure 1A). In the largest subgroup (M1), which consisted of 104 patients (74.2%), most genes were relatively downregulated. In the M2 subgroup, which comprised 38 patients (26.8%), the prognostic genes were frequently overexpressed, with the exception of SCLY and ACSF3. This latter subgroup had a significantly truncated overall survival (OS) in comparison to the M1 subtype (log-rank p -value 0.001; Figure 1B). Similarly, we observed a shorter progression-free survival (PFS) in the M2 subtype ($p = 0.011$; Figure 1C). Interestingly, the clinical features in the two metabolic subtypes were indistinguishable, including features that were known to influence survival (Table 1). Information on diabetes status was available for 122 patients. In patients with the M1 phenotype, 25 of 86 (29.1%) patients had diabetes. In the M2 phenotype, 7 of 34 (20.5%) patients had history of diabetes. This was not significantly different.

A predictive model based on the multilayer perceptron (MLP) classifier was generated for validation studies. The accuracy of this model on internal validation was $95.8\% \pm 0.8\%$. RNASeq data were acquired from three different cohorts to validate our findings. These included the ICGC study ($N = 172$), an analysis of tumor cells enriched by laser capture microdissection (LCM) from surgical patients; the COMPASS trial ($N = 272$) (Aung et al., 2018), LCM-enriched tumor cells from unresectable PDAC (Zhang et al., 2019); and the CPTAC study ($N = 140$), a bulk analysis of mostly surgical patients (Cao et al., 2021). The MLP classifier was based on gene expression z -scores, facilitating the classification of metabolic phenotype, including in bulk transcriptomes. The proportions of each cohort that consisted of the M2 phenotype were approximately the same in each cohort: 23.8% in the ICGC cohort; 23.2% in the COMPASS cohort; and 30.0% in the CPTAC cohort (Supplementary Figure S1). However, in the COMPASS study, a larger proportion of patients with documented metastatic PDAC (54 of 219; 24.7%) were M2 in comparison with the fraction of locally advanced cases that were M2 (2 of 28; 7.1%) ($p = 0.01$).

In each validation cohort, cases with the M2 phenotype had the shortest overall survival (Figures 1D–F). In cases from the COMPASS trial that consisted of patients on palliative chemotherapy, the metabolic subtype also conferred differences in chemotherapy response rates. Specifically, following treatment with FOLFIRINOX (FFX), disease progression was more commonly reported in M2 tumors (48.0% vs. 20.4%; $p = 0.014$), and response rate was higher in M1 tumors. Similarly, progressive disease (PD) was more

common in patients with M2 tumors treated with gemcitabine/nab-paclitaxel (37.5% vs. 12.7%; $p = 0.067$) (Figures 1G, H).

Metabolic subtypes and PDAC molecular subtypes

The metabolic subtypes were evaluated in the context of molecular subtypes described by Collisson et al., 2011; Moffitt et al., 2015; Bailey et al., 2016. Of note, the Moffitt subtypes were also derived from a computational deconvolution method. Using this approach, two tumor phenotypes (classical and basal-like) were identified, and the phenotype of the stromal compartment could also be dichotomized (normal and activated), each with prognostic significance. The relationship of other molecular subtypes with our metabolic subtypes is illustrated in Figures 2A–C. In each case, there were significant differences between M1 and M2 tumors. M2 tumors (with the worst prognosis) were most frequently Moffitt basal-like. In comparison to the M1 subtype, M2 tumors were more frequently Collisson quasimesenchymal and Bailey squamous. The Moffitt basal-like, Collisson quasimesenchymal and Bailey squamous subtypes reportedly have the worst prognosis.

Recently, using pan-cancer data from 33 diverse tumor types, Thorsson et al. identified six immune subtypes based on the composition of the immune infiltrate and expression of immunomodulatory genes (Thorsson et al., 2018). Immunoinflammatory patterns differed significantly in M1 and M2 tumors (p -value < 0.001 ; Figure 2D). Specifically, M1 tumors were comprised of more inflammatory (C3) and TGF- β dominant (C6) immune subtypes; M2 tumors were almost all wound healing (C1) and IFN- γ dominant (C2) subtypes, although 8.1% were also TGF- β dominant subtype. C1 tumors have increased expression of angiogenic genes and a Th2 cell bias in their adaptive immune cell infiltrate. C2 tumors have a strong CD8 signal, but high T cell receptor diversity. In the pan-cancer paper by Thorsson et al., C1 tumors and C2 tumors had a worse prognosis than C3 tumors, and C4 and C6 tumors had the worst survival outcomes.

Biological features of metabolic subtypes

There were no significant differences between M1 and M2 tumors in mutation frequency or copy number variations. To gain an understanding of the biological features that characterized our two metabolic subtypes, we first performed a gene set enrichment analysis (GSEA), focusing first on the 50 hallmark gene sets in MSigDB. Using deconvolved gene expression levels from the tumor compartment from the TCGA dataset, pathways with the most significant enrichment in the poor prognosis M2 subtype included protein secretion, UV response, mammalian target of rapamycin 1 (mTORC1) signaling, and heme metabolism (Table 2). Metabolic pathways that did not quite reach significance included glycolysis, androgen response, angiogenesis and adipogenesis. A similar analysis was performed on the two datasets derived

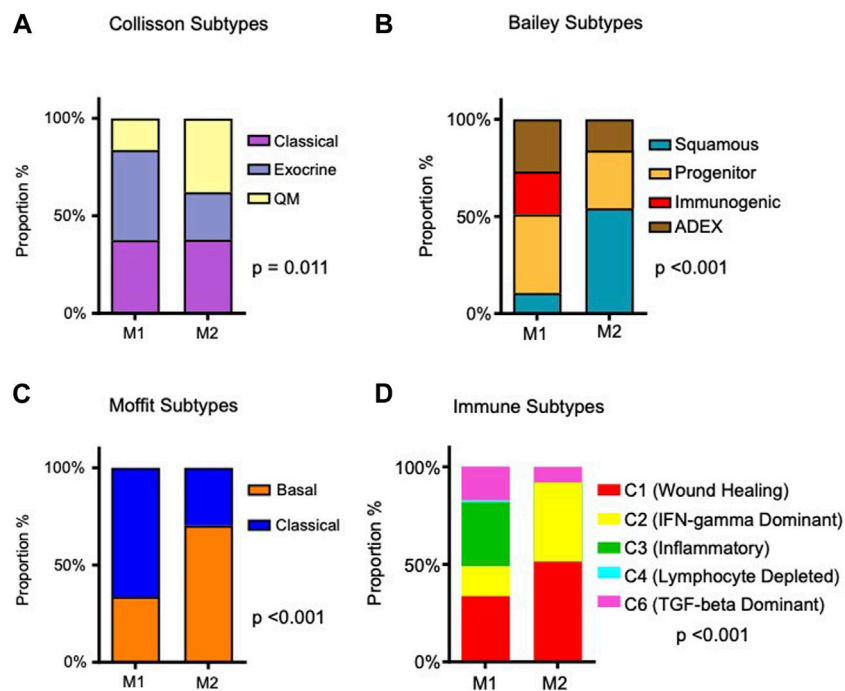


FIGURE 2

Association of identified metabolic subtypes with PDAC molecular subtypes and immune subtypes. (A) Stacked bar charts depicting the distribution of patients in the M1 and M2 PDAC tumor metabolic subtypes, alongside their intersection with previously reported PDAC molecular subtypes (Collisson et al., 2011; Moffit et al., 2015; Bailey et al., 2016). (B) Proportions of immune subtypes as described by Thorsson et al. (2018) are presented within the identified tumor metabolic subtypes. Chi-square *p*-values are provided to assess the significance of associations.

TABLE 2 Gene set enrichment analysis in M1 and M2 subtype in the TCGA cohort (the discovery cohort) using the Broad Institute 50 Hallmarks gene sets.

Enriched Hallmark functions	Normalized enrichment score (NES)	Nominal <i>p</i> -value	FDR (q-value)
Protein secretion*	1.85	0.000	0.214
UV response*	1.62	0.015	0.228
mTORC1 signaling*	1.67	0.022	0.211
Heme metabolism*	1.52	0.039	0.228
Unfolded protein response	1.69	0.020	0.383
DNA repair	1.56	0.022	0.282
MYC targets	1.68	0.035	0.259
Glycolysis	1.45	0.069	0.232
Androgen response	1.47	0.071	0.231
Angiogenesis	1.47	0.083	0.250
Adipogenesis	1.44	0.087	0.229
Peroxisome	1.39	0.097	0.253
G2M checkpoint	1.55	0.105	0.228
E2F targets	1.52	0.127	0.209

* = Significant based on both nominal *p*-value (≤ 0.05) and FDR (≤ 0.25).

from physically enriched tumor cells. In the ICGC dataset, protein secretion, oxidative phosphorylation, MYC targets and DNA repair were enriched. In the COMPASS dataset, protein

secretion and MYC targets were enriched. Finally, using bulk data from CPTAC, protein secretion, MYC targets and DNA repair pathways were enriched.

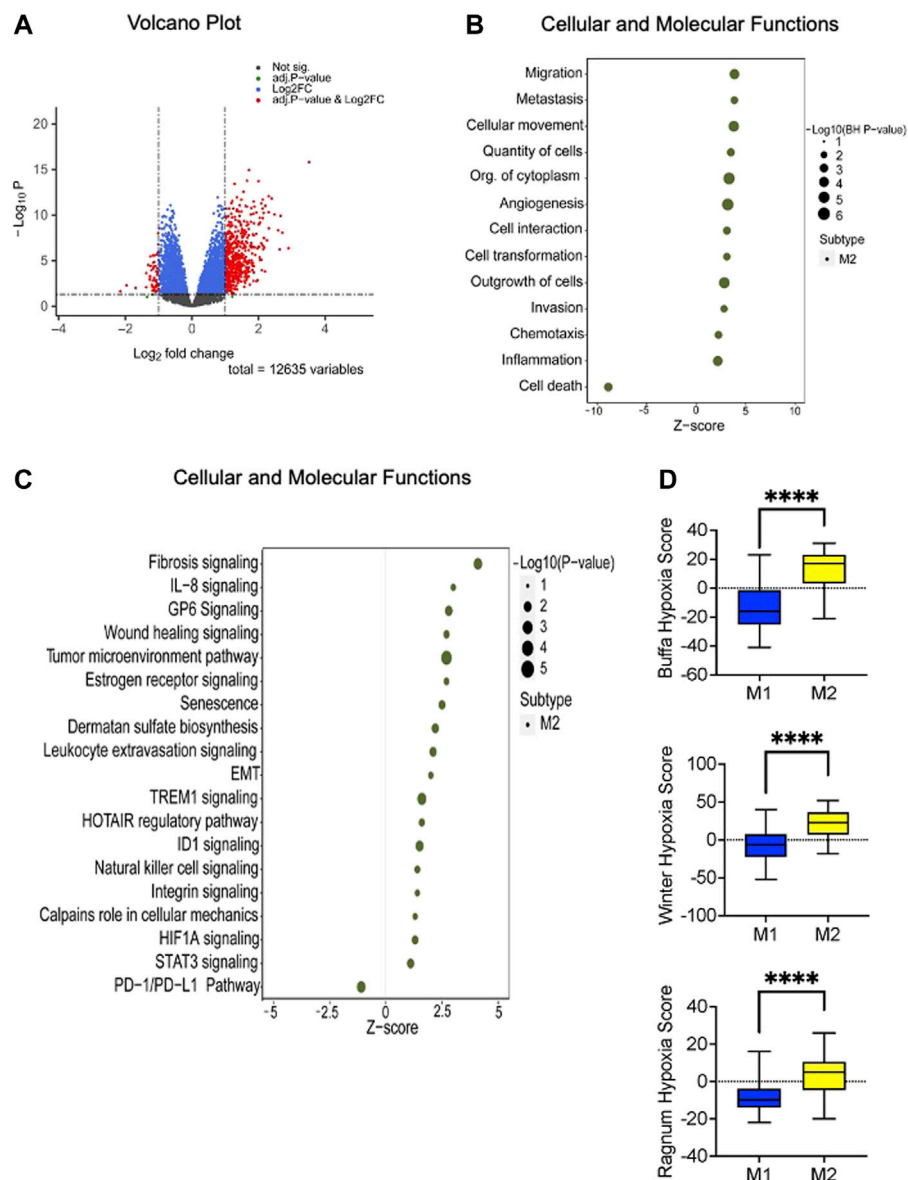


FIGURE 3

Altered canonical pathways and functions linked to the high-risk (M2) subtype. (A) Volcano plot demonstrating the differentially expressed genes ($n = 551$) between the identified metabolic subtypes. Statistical significance is represented by BH adjusted p -value of 0.05 and a \log_2 fold change threshold of ± 1 . (B) Ingenuity Pathway Analysis (IPA) of the differentially expressed genes reveals that the high-risk subtype is positively enriched in cellular proliferation, migration, and invasion. The enrichment analysis yields adjusted p -value less than or equal to 0.05. (C) Disrupted canonical pathways associated with M2 metabolic subtype encompass crucial signaling pathways. Fisher's exact test results indicate a significance level of less than or equal to 0.05. The x-axis represents the activated z-score with sizes indicating the level of significance (right side of the plot). (D) The box plot illustrates the hypoxia scores associated with the identified metabolic subtypes. Based on previous studies (Winter et al., 2007; Buffa et al., 2010; Ragnum et al., 2014) patients with the M2 phenotype exhibits higher scores than M1 phenotype. The statistical significance of these comparisons was assessed using the Mann-Whitney U test, with a p -value threshold set at 0.05 or less.

To derive a deeper understanding of the potential biological features distinguishing the two metabolic subtypes, we applied a knowledge-based approach using Ingenuity Pathways Analysis (Qiagen, Redwood City, CA). There were 551 differentially expressed genes (adjusted p -value 0.05 and \log_2 fold changes equal or greater than ± 1) that informed the analysis (Figure 3A, Supplementary Data Sheet S1). Top molecular and cellular functions positively enriched in M2 tumors included cell migration, invasion, angiogenesis, proliferation, and inflammation (Figure 3B).

Canonical pathways that were significantly enriched included fibrosis signaling, the tumor microenvironment pathway, estrogen receptor signaling, TREM1 signaling, ID1 signaling, STAT3 signaling, and HIF1 signaling (Figure 3C). Interestingly, ID1 expression leads to IL-6 production, activating STAT3 (Meng et al., 2020), which is known to drive tumor associated macrophages toward an immunosuppressive phenotype, also promoting angiogenesis, invasion and epithelial-mesenchymal transition (EMT) (Chang et al., 2013). TREM1 similarly appears to amplify

TABLE 3 Metabolism-focused gene set enrichment analysis of M1 and M2 subtypes.

Enriched metabolic pathways	Normalized enrichment score (NES)	Nominal <i>p</i> -value	FDR (q-value)
Glucagon signaling in metabolic regulation*	2.08	0.000	0.002
Glycogen metabolism*	1.79	0.006	0.094
Regulation of insulin secretion*	1.55	0.039	0.251
Metabolism of water-soluble vitamins and cofactors	1.57	0.017	0.267
Sphingolipid	1.59	0.026	0.321
Phase II conjugation of compounds	1.50	0.031	0.292
Triglyceride metabolism	1.46	0.045	0.312

* = Significant based on both nominal *p*-value (≤ 0.05) and FDR (≤ 0.25).

the inflammatory response, including production of IL-6; blockade of TREM1 has reportedly attenuated tumor growth in an experimental PDAC model (Shen and Sigalov, 2017).

To further dissect the metabolic features that characterized our two metabolic subtypes, we performed a GSEA using only genes with a known metabolic function. The high-risk M2 subtype had significant enrichment in glucagon signaling, glycogen metabolism, and regulation of insulin secretion. Other enriched metabolic pathways (but to a lesser degree) included sphingolipid and triglyceride metabolism (Table 3).

Seemingly, the findings related to carbohydrate metabolism were contradictory. Firstly, in M2 PDACs, glucagon signaling, which promotes gluconeogenesis and glycogen synthesis, was enriched. On the other hand, while the enrichment of glycolysis was not statistically significant (nominal $p = 0.07$, FDR 0.23), virtually all of the key genes corresponding to enzymes promoting glycolysis were elevated. Secondly, there was evidence of both glycogen synthesis (NES 1.50, nominal $p = 0.075$, FDR = 0.054) and glycogenolysis (NES 1.74, nominal $p = 0.009$, FDR = 0.013) (Supplementary Figures S2A, B). The rate limiting enzymes involved in both of these processes are both increased in M2, although there was a 2-fold upregulation of PYGL (promoting glycogenolysis); GYS1 (encouraging glycogen synthesis) was increased by less than 10%. Of note, in cell lines in hypoxic conditions, there is an initial rapid increase in glycogen synthesis catalyzed by GYS1 followed by PYGL-dependent glycogen breakdown (Favaro et al., 2012). The expression of GYS1 is HIF1 α -dependent, but PYGL is not consistently HIF1 α -dependent; PYGL depletion impairs tumor growth, as glycogen breakdown is integral to tumor cell function and growth.

Metabolic phenotype as a function of hypoxia

To explore the potential role of hypoxia, we compared the hypoxia scores between M1 and M2 PDACs using three previously described methods (Winter et al., 2007; Buffa et al., 2010; Ragnum et al., 2014). In this analysis using bulk transcriptomic data, M2 tumors had significantly higher hypoxia scores (Figure 3D). However, when a focused GSEA on the transcriptome ascribed to the tumor compartment was performed using hypoxia-related gene sets obtained from the

Broad Institute MSigDB (Hallmark Hypoxia) and Reactome (Cellular Response to Hypoxia), there was no significant enrichment (nominal p -value 0.11). This suggests that any hypoxia in the tumor microenvironment predominantly affects the stromal compartment rather than the tumor compartment. Other studies have similarly suggested that there is considerable intratumor heterogeneity of hypoxia in human and murine PDAC tumors (Dhani et al., 2015; Lee et al., 2016). Moreover, inflammatory cancer-associated fibroblasts (iCAFs) preferentially localize in hypoxic regions (Mello et al., 2022).

Hypoxia is thought to stem from the desmoplastic stroma of PDAC, possibly by contributing to increased intratumoral pressures, compressing tumor cell vasculature (Chauhan et al., 2014). In that same vein, we evaluated related factors in the two metabolic subtypes. Tumor size was considered, as we postulated that larger tumors would be more susceptible to hypoxia. Neither tumor size nor T-stage were significantly different between M1 and M2 tumors. Angiogenesis is a known consequence of hypoxia (Pugh and Ratcliffe, 2003). Indeed, IPA analysis demonstrated enrichment in genes involved with angiogenesis in M2 tumors. Finally, the proportion of tumor that consists of stroma may also reflect the state of tumor oxygenation. Previously, in PDAC, it was reported that HIF-1 α activation by tumor hypoxia causes secretion of sonic hedgehog (SHH) by cancer cells, which stimulates deposition of fibrous tissue by stromal fibroblasts (Spivak-Kroizman et al., 2013). However, in contrast to that observation, tumor content was significantly greater in the putatively hypoxic M2 tumors, whether estimated histologically (tumor nuclei count), or by ABSOLUTE or DeMixT (Supplementary Figure S3).

Upstream regulators associated with the high-risk metabolic subtype

We took a causal analysis approach using the IPA advanced network tool to infer the identity of upstream regulatory molecules and associated mechanisms related to our two metabolic subtypes (Kr amer et al., 2013). Based on the upstream regulator analysis using the differentially expressed genes, we identified a number of upstream regulators that were activated in the high-risk metabolic subtype, including AGT (z-score 4.115), TGF β 1 (z-score 3.734), SMARCA4 (z-score 3.66), NF κ B (z-score 3.238), STAT1 (z-score 3.106), and VEGF (z-score 3.084). There were also a

number of inhibited upstream regulators associated with the M2 phenotype, including IKZF1 (z score -2.82), IKZF3 (z-score -2.442), EWSR1-FLI1 (z-score -2.236), GMNN (z-score -2.236), and TAF4 (z-score -2.186).

Angiotensinogen (AGT), a substrate for angiotensin II, the primary peptide for the renin-angiotensin system (RAS) was predicted to directly control the expression of 43 DEGs in our dataset: 34 of them were activated, and 9 of them were inhibited (overlap p -value $6.33E-07$, BH p -value $2.72E-05$). Evidence suggested that stimulation of local RAS induces VEGF expression, promoting PDAC tumor growth (Anandanadesan et al., 2007).

TGF β 1 was associated with 62 upregulated DEGs (overlap p -value $4.97E-06$, BH p -value $6.91E-05$), of which 21 overlapped with the AGT network. Angiotensin II (ANG II) increases the expression of TGF β -1 (Weigert et al., 2002). In PDAC, TGF β encourages tumor growth by regulating EMT through SMAD4. It has been previously reported in different types of cancer that cellular proliferation, migration, and angiogenesis could also be influenced by TGF β , partially mediated by VEGF (Li et al., 2005; Ferrari et al., 2009). TGF β is also a prominent mediator of NF κ B activation and PTEN suppression in pancreatic cancer (Chow et al., 2010). SMARCA4 is associated with mitotic spindle, apical junction, and PI3/AKT/mTOR signaling pathways; it is increased in many types of cancer and is reportedly associated with the poor prognosis in other cancers (Peng et al., 2021; Guerrero-Martínez and Reyes, 2018). In all, the aggressive M2 phenotype has a number of potentially causative pathways, some of which may be interrelated.

Single cell studies

While this study described metabolic phenotypes in the tumor compartment (and therefore differed from bulk transcriptomic studies), in effect the deconvolved gene expression data that comprised the M1 and M2 phenotypes are the product of the average gene expression values in tumor cells. The question remained whether tumor cells existed in a uniform metabolic state within a tumor, or whether there was phenotypic diversity at the single cell level. To explore this, we accessed data from a single cell RNA-Seq studies on 16 tumors (10 primary tumors and 6 metastases) (Lin, W. et al., 2020). Epithelial tumor cells were identified, and expression levels for each of the metabolic genes associated with PDAC prognosis were determined for each individual tumor cell. Samples with at least 100 epithelial tumor cells (ETCs) were used in this analysis. One metastasis sample (MET03) and three primary tumors (P01, P02, and P03) were excluded from analysis because of insufficient numbers of cells. [Supplementary Figure S4](#) shows the degree of heterogeneity within each tumor. For ease of comparison, the gene order on the heatmap is the same as in [Figure 1A](#).

There were several interesting observations. First, there was significant diversity in the expression of prognostic metabolic genes between cells within any individual tumor. Second, the proportion of tumor cells with greater similarity to M1 or M2 phenotype varied little. Third, the genes that were characteristically upregulated in M2 PDACs were not as closely co-related at the single cell level. Finally, there were individual

tumors where more of the prognostic metabolic genes were upregulated in more cells. Likely, these types of tumors represented PDACs that would be recognizable as M2 PDACs in the bulk analysis.

Potential drug targets

We used the PharmacGx platform to determine some potential drug targets that might particularly target the high-risk M2 subtype. Initially, we obtained two different PDAC cell line datasets including the Cancer Cell Line Encyclopedia (CCLE, $n = 41$) (Cancer Cell Line Encyclopedia Consortium, Genomics of Drug Sensitivity in Cancer Consortium, 2015), and the Genetech Cell Line Screening Initiative (gCSI, $n = 35$) (Klijn et al., 2014). The multilayer perceptron classifier was applied to cell lines to classify metabolic phenotype. Of the CCLE PDAC cell lines, 9 (22.0%) had the M2 phenotype; 10 of the gCSI cell lines (28.6%) were M2 ([Supplementary Figure S5](#)). In the CCLE collection, there was no differential drug sensitivity between M1 and M2 cell lines. In gCSI cell lines, based on the nominal p -value, M2 cell lines were more sensitive to docetaxel and erlotinib ($p = 0.048$ and $p = 0.019$, respectively). However, following correction for multiple comparisons, this was not significant ([Supplementary Figure S6](#)).

Discussion

Metabolic reprogramming is one of the hallmarks of cancer (Hanahan and Weinberg, 2011). As with other hallmarks of cancer, metabolic reprogramming confers a growth advantage to malignant cells. Metabolic derangements are potentially attractive therapeutic targets, especially perturbations that adversely affect patient health and survival. We have identified two novel metabolic subtypes based on metabolic genes most closely associated with survival events. One of these phenotypes (M2) has a significantly worse prognosis. The M2 phenotype has a significant overlap with the poor prognosis molecular subtypes described by Collisson et al., 2011; Moffitt et al., 2015; Bailey et al., 2016. The pattern of immune infiltrate or M2 tumors differs from M1 tumors. In addition to biological features that would encourage tumor growth, M2 PDACs appear to be hypoxic (in the stromal compartment) and have increased glycogen turnover.

Others have described metabolic subtypes of PDAC (Daemen et al., 2015; Peng et al., 2018; Karasinska et al., 2020). The approach taken by others has involved identifying tumors based on selective metabolic pathway genes. Our approach was different. Instead of evaluating known metabolic pathways as a whole, we took an unbiased approach to identify metabolic genes that most consistently associated with survival events. The result was a list of 26 genes with diverse functions; their pattern of expression dictated one of two metabolic phenotypes with distinct biological and clinical features.

One challenge in the study of PDAC is its cellular heterogeneity. PDACs vary considerably in their composition, some containing large numbers of tumor cells and others dominated by desmoplastic stroma. The stromal constituents are similarly diverse. As a result, there are limitations in how one may interpret molecular studies

from bulk samples. To address this, we used a computational deconvolution method to determine the proportion of expression of each gene that could be ascribed to the tumor compartment. A simple two-compartment model was applied using DeMixT. There are numerous other approaches to computational deconvolution, many capable of deconvolution to many cell types (Newman et al., 2015; Avila Cobos et al., 2020). Technical and biological biases can affect the accuracy of any deconvolution technique (Vallania et al., 2018). The metabolic subtypes we described in deconvolved data were observable in samples derived from physical enrichment, by laser capture microdissection (LCM). LCM too has limitations: it is virtually impossible to isolate an entirely pure population of tumor cells, and thermal injury to cells may alter any gene expression measurements. However, it was reassuring that the model built from computational deconvolution could be replicated in physically enriched tumor cells.

Even in a pure tumor cell population isolated from a PDAC, it is likely that there is significant heterogeneity in cells from the same tumor. Therefore, the metabolic features attributable to our two metabolic subtypes represent an average of what may be present throughout the entire tumor sample. Indeed, using sc-RNASeq data, we found that there is substantial cell-to-cell diversity in the expression levels of the individual genes within any single PDAC. Further studies will be required using high resolution spatial transcriptomics methods to identify the determinants of intratumoral heterogeneity in biological processes such as metabolism. However, current methods may not have sufficient resolution at the cellular level, and gene representation of any single cell analysis technique is still not comprehensive.

Unlike previous studies (Daemen et al., 2015; Peng et al., 2018), we did not specifically evaluate any drugs targeting metabolic pathways. Rather, we evaluated the clinical and *in vitro* responses to chemotherapies currently in clinical use. In the COMPASS cohort, M1 tumors had a higher rate of PR and SD in response to FOLFIRINOX; a trend toward higher response rate was seen in M1 tumors treated with gemcitabine + paclitaxel. *In vitro*, M2 tumors were slightly more sensitive to docetaxel, another taxane. We view the *in vitro* data as weak, as only one of the cell line collections demonstrated a slight difference in sensitivity, and this could not be replicated in the other cell line collection. Another reason for this discrepancy may be the inclusion of a fluoropyrimidine (5-FU) in FOLFIRINOX and gemcitabine in paclitaxel-gemcitabine; high DPYD expression in M2 tumors confers resistance to fluoropyrimidines and gemcitabine (Gustavsson et al., 2009; Zhang et al., 2017; Tsukahara et al., 2022). Our *in silico* experiments were limited by statistical power. Further studies would be required to explore the utility of our metabolic classification in treatment selection, especially as it pertains to drugs interfering with specific metabolic functions.

Perhaps the most well described feature of cancer cells is their high rate of glycolysis followed by lactic acid fermentation even in normoxic conditions, known as the Warburg effect. One particularly intriguing feature of M2 tumors was the simultaneous upregulation of genes related to glycogen synthesis, glycogenolysis, and glycolysis. Under nutrient deprivation, cancer cells exploit glycogen metabolism for optimal glucose utilization. Different *in vitro*

studies have demonstrated glycogen accumulation without inhibiting or interfering with the breakdown process (Favaro et al., 2012; Pelletier et al., 2012).

Glycogen accumulation and breakdown have been variably observed under hypoxic conditions. The relative rates of those opposing pathways changes with the chronicity of hypoxia, where there is early glycogen accumulation followed by a gradual decline (Favaro et al., 2012). GYS1 induction (and therefore glycogen accumulation) is dependent on HIF1 α , a central factor in the response to hypoxia. Indeed, M2 tumors may be a result of a hypoxic microenvironment. Using previously validated gene panels on bulk samples, hypoxia scores were higher in M2 tumors. Consistent with this, angiogenesis-related genes were upregulated. On the other hand, specifically in the tumor compartment, only 32 of 212 hypoxia related genes were dysregulated. This suggests that changes in hypoxia-related genes mostly resided in the stromal compartment, which is consistent with previous report (Mello et al., 2022).

Finally, the potential role of the renin-angiotensin in M2 tumors was interesting. The local RAS system encourages tumor cell proliferation and angiogenesis by upregulating EGFR and VEGF expression (Anandanadesan et al., 2007) and RAS inhibition induces apoptosis in pancreatic cancer cells (Gong et al., 2010). In a murine model of PDAC, gemcitabine and the angiotensin receptor blocker (ARB) losartan synergistically inhibit tumor growth via VEGF suppression (Noguchi et al., 2009). There have been clinical reports of survival benefits in PDAC patients receiving RAS inhibitors who were treated with gemcitabine (Nakai et al., 2010) and with resection (Cerullo et al., 2017). A large retrospective study of 8,158 PDAC patients reported survival benefits related to RAS inhibition, more pronounced with ARBs than angiotensin I converting enzyme (ACE) inhibitors (Keith et al., 2022). Further studies will be required to determine whether our metabolic subtypes confer differential sensitivity to RAS inhibition.

In conclusion, we have described two metabolic subtypes based on gene expression in tumor cells. Our unbiased approach in identifying the clinically most impactful genes unveiled novel phenotypes. We are unable to discern whether there are host factors that contribute to the metabolic phenotypes. The incidence of diabetes was similar in patients with M1 and M2 tumors, but BMI data were lacking. Race distribution was similar in patients with M1 and M2 tumors, but the majority of patients were white. Further studies in patients who are not white are warranted. We have not yet identified how our tumor classification is relevant to therapeutic decisions. Indeed, none of the molecular subtypes described for PDAC have yet proven relevant to clinical decision-making. However, some interesting biological phenomena have been uncovered that warrant further investigation. Geospatial studies would be particularly informative in this regard.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: <https://portal.gdc.cancer.gov>, <https://dcc.icgc.org>, <https://ega-archive.org/studies/EGAS00001002543>, <https://www>.

orcestra.ca, <https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE154778>.

Ethics statement

The studies involving humans were approved by the Enrichment of Genomics with Clinically Informative Molecular Phenotyping (Personalized OncoMetabolomics). The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

JP: writing—original draft, conceptualization, data curation, formal analysis, methodology, validation, visualization, Investigation. MA: Data curation, Investigation, Writing—review and editing. SC: Data curation, Investigation, Writing—review and editing. GJ: Data curation, Investigation, Writing—review and editing. NF: Formal analysis, Investigation, Writing—review and editing. BH-K: Formal analysis, Investigation, Writing—review and editing. JK: Investigation, Writing—review and editing. GO'K: Data curation, Investigation, Writing—review and editing. SG: Data curation, Funding acquisition, Investigation, Writing—review and editing. DS: Funding acquisition, Investigation, Writing—review and editing. DR: Funding acquisition, Investigation, Writing—review and editing. GZ: Funding acquisition, Investigation, Writing—review and editing. OB: Conceptualization, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Writing—original draft, Writing—review and editing.

Funding

The authors declare financial support was received for the research, authorship, and/or publication of this article. This research was supported through philanthropic donations, and funding from the Terry Fox Research Institute.

Acknowledgments

We would like to thank Farshad Farshidfar and Cynthia Stretch for their support and providing necessary resources for this study.

Conflict of interest

HighLifeR is owned by Qualisure Diagnostics Inc., and OB is director and shareholder of Qualisure Diagnostics Inc.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fgene.2023.1282824/full#supplementary-material>

SUPPLEMENTARY FIGURE S1

Visualization of predicted metabolic subtypes in all validation cohorts. The heatmaps exhibit the expression of 26 prognostic metabolic genes across the identified metabolic subtypes within each validation cohort. The proportion of patients belong to each subgroup are shown in the figures. The gene order follows the same order as depicted in the discovery cohort. Expression of identified prognostic metabolic genes are similar to the discovery set and highlight noticeable differences between groups.

SUPPLEMENTARY FIGURE S2

Glycogen metabolism related gene expression in identified metabolic subtypes. (A) Gene set enrichment analysis focussed on glycogen metabolism, showing enrichment of glycogen catabolism in the M2 phenotype accompanied by enrichment of glycogen synthesis pathways. (B) Scatter plot shows expression levels of regulators involved in glycogen synthesis and degradation process. Mann-Whitney U tests p values are shown at the top of the plots.

SUPPLEMENTARY FIGURE S3

Tumor cell composition in metabolic subtypes. Three different methods were applied to estimate the proportion of tumor that is comprised of tumor cells: Absolute purity, DeMixT and tumor nuclei counts. Violin plots show higher tumor content in the M2 phenotype and the differences in tumor content are significantly different between M1 and M2 based on p-value <0.05.

SUPPLEMENTARY FIGURE S4

Expression of prognostic metabolic genes at a single cell level. Single cell data were used to determine the expression of prognostic genes in malignant epithelial cells. Each heatmap represents gene expression of prognostic metabolic genes in a single tumor. The x-axis depicts data from individual epithelial tumor cells, and the y-axis depicts data from the 26 prognostic metabolic genes.

SUPPLEMENTARY FIGURE S5

Predicted metabolic phenotypes in PDAC cell lines. The heatmap shows two distinct groups based on the 26 prognostic metabolic gene expressions in individual PDAC cell line cohorts (CCLE and gCSI). In each cell line collection, the proportion of cells with each metabolic subtype are similar.

SUPPLEMENTARY FIGURE S6

Drug target analysis of PDAC cell lines. Sensitivity to antineoplastic drugs was assessed in cell lines identified as M1 and M2 cell lines. Each box plot shows area above the drug-response curve (AAC) values for drug screening in CCLE and gCSI cohorts, respectively. The p-values based on the Wilcoxon test are indicated on the top of box plots.

SUPPLEMENTARY TABLE S1

Prognostic metabolic genes, and summary of functions.

SUPPLEMENTARY DATA SHEET S1

Prognostic metabolic genes and differentially expressed genes list.

References

- Avila Cobos, F., Alquicira-Hernandez, J., Powell, J. E., Mestdagh, P., and De Preter, K. (2020). Benchmarking of cell type deconvolution pipelines for transcriptomics data. *Nat. Commun.* 11 (1), 5650. doi:10.1038/s41467-020-19015-1
- Alpertunga, I., Sadiq, R., Pandya, D., Lo, T., Dulgher, M., Evans, S., et al. (2021). Glycemic control as an early prognostic marker in advanced pancreatic cancer. *Front. Oncol.* 11, 571855. doi:10.3389/fonc.2021.571855
- Anandanadesan, R., Gong, Q., Chipitsyna, G., Witkiewicz, A., Yeo, C. J., and Arafat, H. A. (2007). Angiotensin II induces vascular endothelial growth factor in pancreatic cancer cells through an angiotensin II type 1 receptor and ERK1/2 signaling. *J. Gastrointest. Surg.* 12 (1), 57–66. doi:10.1007/s11605-007-0403-9
- Aung, K. L., Fischer, S. E., Denroche, R. E., Jang, G. H., Dodd, A., Creighton, S., et al. (2018). Genomics-driven precision medicine for advanced pancreatic cancer: early results from the compass trial. *Clin. Cancer Res.* 24 (6), 1344–1354. doi:10.1158/1078-0432.ccr-17-2994
- Bailey, P., Chang, D. K., Nones, K., Johns, A. L., Patch, A. M., Gingras, M. C., et al. (2016). Genomic analyses identify molecular subtypes of pancreatic cancer. *Nature* 531 (7592), 47–52. doi:10.1038/nature16965
- Bathe, O. F., Shaykhutdinov, R., Kopciuk, K., Weljie, A. M., McKay, A., Sutherland, F. R., et al. (2011). Feasibility of identifying pancreatic cancer based on serum metabolomics. *Cancer Epidemiol. Biomarkers Prev.* 20 (1), 140–147. doi:10.1158/1055-9965.epi-10-0712
- Buffa, F. M., Harris, A. L., West, C. M., and Miller, C. J. (2010). Large meta-analysis of multiple cancers reveals a common, compact and highly prognostic hypoxia metagene. *Br. J. Cancer* 102 (2), 428–435. doi:10.1038/sj.bjc.6605450
- Cancer Cell Line Encyclopedia Consortium; Genomics of Drug Sensitivity in Cancer Consortium (2015). Pharmacogenomic agreement between two cancer cell line data sets. *Nature* 528 (7580), 84–87. doi:10.1038/nature15736
- Cao, L., Huang, C., Cui Zhou, D., Hu, Y., Lih, T. M., Savage, S. R., et al. (2021). Proteomic characterization of pancreatic ductal adenocarcinoma. *Cell* 184 (19), 5031–5052.e26. doi:10.1016/j.cell.2021.08.023
- Cao, S., Wang, J. R., Ji, S., Yang, P., Dai, Y., Guo, S., et al. (2022). Estimation of tumor cell total mrna expression in 15 cancer types predicts disease progression. *Nat. Biotechnol.* 40 (11), 1624–1633. doi:10.1038/s41587-022-01342-x
- Cerullo, M., Gani, F., Chen, S. Y., Canner, J. K., and Pawlik, T. M. (2017). Impact of angiotensin receptor blocker use on overall survival among patients undergoing resection for pancreatic cancer. *World J. Surg.* 41 (9), 2361–2370. doi:10.1007/s00268-017-4021-8
- Chang, Q., Bournazou, E., Sansone, P., Berishaj, M., Gao, S. P., Daly, L., et al. (2013). The IL-6/jak/STAT3 feed-forward loop drives tumorigenesis and Metastasis. *Neoplasia* 15 (7), 848–862. doi:10.1593/neo.13706
- Chauhan, V. P., Boucher, Y., Ferrone, C. R., Roberge, S., Martin, J. D., Stylianopoulos, T., et al. (2014). Compression of pancreatic tumor blood vessels by hyaluronan is caused by solid stress and not interstitial fluid pressure. *Cancer Cell* 26 (1), 14–15. doi:10.1016/j.ccr.2014.06.003
- Chow, J. Y., Ban, M., Wu, H. L., Nguyen, F., Huang, M., Chung, H., et al. (2010). TGF-beta downregulates PTEN via activation of NF-kappaB in pancreatic cancer cells. *Am. J. Physiology-Gastrointestinal Liver Physiology* 298 (2), G275–G282. doi:10.1152/ajpgi.00344.2009
- Collisson, E. A., Sadanandam, A., Olson, P., Gibb, W. J., Truitt, M., Gu, S., et al. (2011). Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. *Nat. Med.* 17 (4), 500–503. doi:10.1038/nm.2344
- Craig, S., (2022). A clinically useful and biologically informative genomic classifier for papillary thyroid cancer[Preprint]. <https://www.biorxiv.org/content/10.1101/2022.12.27.522039v1>.
- Daemen, A., Peterson, D., Sahu, N., McCord, R., Du, X., Liu, B., et al. (2015). Metabolite profiling stratifies pancreatic ductal adenocarcinomas into subtypes with distinct sensitivities to metabolic inhibitors. *Proc. Natl. Acad. Sci.* 112 (32), E4410–E4417. doi:10.1073/pnas.1501605112
- Dhani, N. C., Serra, S., Pintilie, M., Schwock, J., Xu, J., Gallinger, S., et al. (2015). Analysis of the intra- and intertumoral heterogeneity of hypoxia in pancreatic cancer patients receiving the nitroimidazole Tracer Pimonidazole. *Br. J. Cancer* 113 (6), 864–871. doi:10.1038/bjc.2015.284
- Fang, Y., Pei, S., Huang, K., Xu, F., Xiang, G., et al. (2022). Single-cell transcriptome reveals the metabolic and clinical features of a highly malignant cell subpopulation in pancreatic ductal adenocarcinoma. *Front. Cell Dev. Biol.* 10, 798165. doi:10.3389/fcell.2022.798165
- Favaro, E., Bensaad, K., Chong, M. G., Tennant, D. A., Ferguson, D. J. P., Snell, C., et al. (2012). Glucose utilization via glycogen phosphorylase sustains proliferation and prevents premature senescence in cancer cells. *Cell Metab.* 16 (6), 751–764. doi:10.1016/j.cmet.2012.10.017
- Ferrari, G., Cook, B. D., Terushkin, V., Pintucci, G., and Mignatti, P. (2009). Transforming growth factor-beta 1 (TGF-beta1) induces angiogenesis through vascular endothelial growth factor (VEGF)-mediated apoptosis. *J. Cell. Physiology* 219 (2), 449–458. doi:10.1002/jcp.21706
- Gong, Q., Davis, M., Chipitsyna, G., Yeo, C. J., and Arafat, H. A. (2010). Blocking angiotensin II type 1 receptor triggers apoptotic cell death in human pancreatic cancer cells. *Pancreas* 39 (5), 581–594. doi:10.1097/mpa.0b013e318c314cd
- Guerrero-Martínez, J. A., and Reyes, J. C. (2018). High expression of SMARCA4 or SMARCA2 is frequently associated with an opposite prognosis in cancer. *Sci. Rep.* 8 (1), 2043. doi:10.1038/s41598-018-20217-3
- Gustavsson, B., Kaiser, C., Carlsson, G., Wettergren, Y., Odin, E., Lindskog, E. B., et al. (2009). Molecular determinants of efficacy for 5-FU-based treatments in advanced colorectal cancer: mrna expression for 18 chemotherapy-related genes. *Int. J. Cancer* 124 (5), 1220–1226. doi:10.1002/ijc.23852
- Hanahan, D., and Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell* 144 (5), 646–674. doi:10.1016/j.cell.2011.02.013
- Jeon, C. Y., Kim, S., Lin, Y. C., Risch, H. A., Goodarzi, M. O., Nuckols, T. K., et al. (2021). Prediction of pancreatic cancer in diabetes patients with worsening glycemic control. *Cancer Epidemiol. Biomarkers Prev.* 31 (1), 242–253. doi:10.1158/1055-9965.epi-21-0712
- Karasińska, J. M., Topham, J. T., Kalloger, S. E., Jang, G. H., Denroche, R. E., Culibrk, L., et al. (2021). Altered gene expression along the glycolysis–cholesterol synthesis axis is associated with outcome in pancreatic cancer. *Clin. Cancer Res.* 26 (1), 135–146. doi:10.1158/1078-0432.ccr-19-1543
- Keith, S. W., Maio, V., Arafat, H. A., Alcusky, M., Karagiannis, T., Rabinowitz, C., et al. (2022). Angiotensin blockade therapy and survival in pancreatic cancer: a population study. *BMC Cancer* 22 (1), 150. doi:10.1186/s12885-022-09200-4
- Klijn, C., Durinck, S., Stawiski, E. W., Haverty, P. M., Jiang, Z., Liu, H., et al. (2014). A comprehensive transcriptional portrait of human cancer cell lines. *Nat. Biotechnol.* 33 (3), 306–312. doi:10.1038/nbt.3080
- Krämer, A., Green, J., Pollard, J., and Tugendreich, S. (2013). Causal analysis approaches in ingenuity pathway analysis. *Bioinformatics* 30 (4), 523–530. doi:10.1093/bioinformatics/btt703
- Law, C. W., Chen, Y., Shi, W., and Smyth, G. K. (2014). Voom: precision weights unlock linear model analysis tools for RNA-seq read counts. *Genome Biol.* 15 (2), R29. doi:10.1186/gb-2014-15-2-r29
- Lee, K. E., Spata, M., Bayne, L. J., Buza, E. L., Durham, A. C., Allman, D., et al. (2016). *hifa1a* deletion reveals pro-neoplastic function of B cells in pancreatic neoplasia. *Cancer Discov.* 6 (3), 256–269. doi:10.1158/2159-8290.cd-15-0822
- Li, H., and Gui, J. (2004). Partial Cox regression analysis for high-dimensional microarray gene expression data. *Bioinformatics* 20 (1), i208–i215. doi:10.1093/bioinformatics/bth900
- Li, Z.-D., Bork, J. P., Krueger, B., Patsenker, E., Schulze-Krebs, A., Hahn, E. G., et al. (2005). VEGF induces proliferation, migration, and TGF-beta1 expression in mouse glomerular endothelial cells via mitogen-activated protein kinase and phosphatidylinositol 3-kinase. *Biochem. Biophysical Res. Commun.* 334 (4), 1049–1060. doi:10.1016/j.bbrc.2005.07.005
- Lin, W., Noel, P., Borazanci, E. H., Lee, J., Amini, A., Han, I. W., et al. (2020). Single-cell transcriptome analysis of tumor and stromal compartments of pancreatic ductal adenocarcinoma primary tumors and metastatic lesions. *Genome Med.* 12 (1), 80. doi:10.1186/s13073-020-00776-9
- Mayers, J. R., Wu, C., Clish, C. B., Kraft, P., Torrence, M. E., Fiske, B. P., et al. (2014). Elevation of circulating branched-chain amino acids is an early event in human pancreatic adenocarcinoma development. *Nat. Med.* 20 (10), 1193–1198. doi:10.1038/nm.3686
- Mello, A. M., Ngodup, T., Lee, Y., Donahue, K. L., Li, J., Rao, A., et al. (2022). Hypoxia promotes an inflammatory phenotype of fibroblasts in pancreatic cancer. *Oncogenesis* 11 (1), 56. doi:10.1038/s41389-022-00434-2
- Meng, J., Liu, K., Shao, Y., Feng, X., Ji, Z., Chang, B., et al. (2020). ID1 confers cancer cell chemoresistance through STAT3/ATF6-mediated induction of autophagy. *Cell Death Dis.* 11 (2), 137. doi:10.1038/s41419-020-2327-1
- Moffitt, R. A., Marayati, R., Flate, E. L., Volmar, K. E., Loeza, S. G. H., Hoadley, K. A., et al. (2015). Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. *Nat. Genet.* 47 (10), 1168–1178. doi:10.1038/ng.3398
- Murtagh, F. (1991). Multilayer perceptrons for classification and regression. *Neurocomputing* 2 (5–6), 183–197. doi:10.1016/0925-2312(91)90023-5
- Nakai, Y., Isayama, H., Ijichi, H., Sasaki, T., Sasahira, N., Hirano, K., et al. (2010). Inhibition of renin–angiotensin system affects prognosis of advanced pancreatic cancer receiving Gemcitabine. *Br. J. Cancer* 103 (11), 1644–1648. doi:10.1038/sj.bjc.6605955
- Newman, A. M., Liu, C. L., Green, M. R., Gentles, A. J., Feng, W., Xu, Y., et al. (2015). Robust enumeration of cell subsets from tissue expression profiles. *Nat. Methods* 12 (5), 453–457. doi:10.1038/nmeth.3337
- Noguchi, R., Yoshiji, H., Ikenaka, Y., Namisaki, T., Kitade, M., Kaji, K., et al. (2009). Synergistic inhibitory effect of gemcitabine and angiotensin type-1 receptor blocker, losartan, on murine pancreatic tumor growth via anti-angiogenic activities. *Oncol. Rep.* 22 (2), 355–360. doi:10.3892/or_00000445

- Pelletier, J., Bellot, G., Gounon, P., Lacas-Gervais, S., Pouysségur, J., and Mazure, N. M. (2012). Glycogen synthesis is induced in hypoxia by the hypoxia-inducible factor and promotes cancer cell survival. *Front. Oncol.* 2, 18. doi:10.3389/fonc.2012.00018
- Peng, L., Li, J., Wu, J., Xu, B., Wang, Z., Giamas, G., et al. (2021). A pan-cancer analysis of SMARCA4 alterations in human cancers. *Front. Immunol.* 12, 762598. doi:10.3389/fimmu.2021.762598
- Peng, X., Chen, Z., Farshidfar, F., Xu, X., Lorenzi, P. L., Wang, Y., et al. (2018). Molecular characterization and clinical relevance of metabolic expression subtypes in human cancers. *Cell Rep.* 23 (1), 255–269.e4. doi:10.1016/j.celrep.2018.03.077
- Pugh, C. W., and Ratcliffe, P. J. (2003). Regulation of angiogenesis by hypoxia: role of the HIF system. *Nat. Med.* 9 (6), 677–684. doi:10.1038/nm0603-677
- Ragnum, H. B., Vlatkovic, L., Lie, A. K., Axcrone, K., Julin, C. H., Friksstad, K. M., et al. (2014). The tumour hypoxia marker pimonidazole reflects a transcriptional programme associated with aggressive prostate cancer. *Br. J. Cancer* 112 (2), 382–390. doi:10.1038/bjc.2014.604
- Raphael, B. J., and Cancer Genome Atlas Research Network (2017). Integrated genomic characterization of pancreatic ductal adenocarcinoma. *Cancer Cell* 32 (2), 185–203.e13. doi:10.1016/j.ccell.2017.07.007
- Rupert, J. E., Narasimhan, A., Jengelly, D. H. A., Jiang, Y., Liu, J., Au, E., et al. (2021). Tumor-derived IL-6 and trans-signaling among tumor, fat, and muscle mediate pancreatic cancer cachexia. *J. Exp. Med.* 218 (6), e20190450. doi:10.1084/jem.20190450
- Shen, Z. T., and Sigalov, A. B. (2017). Novel Trem-1 inhibitors attenuate tumor growth and prolong survival in experimental pancreatic cancer. *Mol. Pharm.* 14 (12), 4572–4582. doi:10.1021/acs.molpharmaceut.7b00711
- Spivak-Kroizman, T. R., Hostetter, G., Posner, R., Aziz, M., Hu, C., Demeure, M. J., et al. (2013). Hypoxia triggers hedgehog-mediated tumor–stromal interactions in pancreatic cancer. *Cancer Res.* 73 (11), 3235–3247. doi:10.1158/0008-5472.can-11-1433
- Thorsson, V., Gibbs, D. L., Brown, S. D., Wolf, D., Bortone, D. S., Ou Yang, T. H., et al. (2018). The immune landscape of cancer. *Immunity* 48 (4), 812–830.e14. doi:10.1016/j.immuni.2018.03.023
- Tsukahara, S., Shiota, M., Takamatsu, D., Nagakawa, S., Matsumoto, T., Kiyokoba, R., et al. (2022). Cancer genomic profiling identified dihydropyrimidine dehydrogenase deficiency in bladder cancer promotes sensitivity to gemcitabine. *Sci. Rep.* 12 (1), 8535. doi:10.1038/s41598-022-12528-3
- Vallania, F., Tam, A., Lofgren, S., Schaffert, S., Azad, T. D., Bongen, E., et al. (2018). Leveraging heterogeneity across multiple datasets increases cell-mixture deconvolution accuracy and reduces biological and technical biases. *Nat. Commun.* 9 (1), 4735. doi:10.1038/s41467-018-07242-6
- Wang, Yu, Liang, Y., Xu, H., Zhang, X., Mao, T., Cui, J., et al. (2021). Single-cell analysis of pancreatic ductal adenocarcinoma identifies a novel fibroblast subtype associated with poor prognosis but better immunotherapy response. *Cell Discov.* 7 (1), 36. doi:10.1038/s41421-021-00271-4
- Wang, Z., Cao, S., Morris, J. S., Ahn, J., Liu, R., Tyekucheva, S., et al. (2018). Transcriptome deconvolution of heterogeneous tumor samples with immune infiltration. *iScience* 9, 451–460. doi:10.1016/j.isci.2018.10.028
- Weigert, C., Brodbeck, K., Klopfer, K., Häring, H. U., and Schleicher, E. D. (2002). Angiotensin II induces human TGF- β 1 promoter activation: similarity to hyperglycaemia. *Diabetologia* 45 (6), 890–898. doi:10.1007/s00125-002-0843-4
- Winter, S. C., Buffa, F. M., Silva, P., Miller, C., Valentine, H. R., Turley, H., et al. (2007). Relation of a hypoxia metagene derived from head and neck cancer to prognosis of multiple cancers. *Cancer Res.* 67 (7), 3441–3449. doi:10.1158/0008-5472.can-06-3322
- Witten, I. H., (2017). *Data mining: practical machine learning tools and techniques*. Cambridge, MA, USA: Morgan Kaufmann.
- Zhang, C., Liu, H., Ma, B., Song, Y., Gao, P., Xu, Y., et al. (2017). The impact of the expression level of intratumoral dihydropyrimidine dehydrogenase on chemotherapy sensitivity and survival of patients in gastric cancer: a meta-analysis. *Dis. Markers* 2017, 9202676–9202711. doi:10.1155/2017/9202676
- Zhang, J., Bajari, R., Andric, D., Gerthoffert, F., Lepsa, A., Nahal-Bose, H., et al. (2019). The international cancer Genome Consortium data portal. *Nat. Biotechnol.* 37 (4), 367–369. doi:10.1038/s41587-019-0055-9
- Zhong, X., Narasimhan, A., Silverman, L. M., Young, A. R., Shahda, S., Liu, S., et al. (2022). Sex specificity of pancreatic cancer cachexia phenotypes, mechanisms, and treatment in mice and humans: role of activin. *J. Cachexia, Sarcopenia Muscle* 13 (4), 2146–2161. doi:10.1002/jcsm.12998