Article

Not peer-reviewed version

Molecular Pathway and Immune Profile Analysis of IPMN-Derived versus PanIN-Derived Pancreatic Ductal Adenocarcinomas

Margaret A Park , Kristyn Gumpper-Fedus , Somashekar G. Krishna , Maria C. Genilo-Delgado , Stephen Brantley , Phil A. Hart , Mary E. Dillhoff , Maria F. Gomez , Toni L. Basinski , Shaffer R. Mok , Anjuli K. Luthra , Jason B. Fleming , Amir Mohammadi , Barbara A Centeno , Kun Jiang , Alexandra Karolak , Daniel Jeong , Dung-Tsa Chen , Paul Stewart , Jamie K. Teer , Zobeida Cruz-Monserrate * , Jennifer B. Permuth *

Posted Date: 26 September 2024

doi: 10.20944/preprints202409.2045.v1

Keywords: pancreatic ductal adenocarcinoma; IPMN; PanIN; transcriptomics

Article

Molecular Pathway and Immune Profile Analysis of IPMN-Derived versus PanIN-Derived Pancreatic Ductal Adenocarcinomas

Margaret A. Park [1,2], Kristyn Gumpper-Fedus [3], Somashekar G. Krishna [3], Maria C. Genilo-Delgado [1], Stephen Brantley [4], Phil A. Hart [3], Mary E. Dillhoff [5], Maria F. Gomez [1,6], Toni L. Basinski [1], Shaffer R. Mok [1], Anjuli K. Luthra [1], Jason B. Fleming [7], Amir Mohammadi [1], Barbara A. Centeno [4], Kun Jiang [4], Aleksandra Karolak [8], Daniel Jeong [9], Dung-Tsa Chen [2], Paul A. Stewart [2], Jamie K. Teer [2,†], Zobeida Cruz-Monserrate [2,*,†] and Jennifer B. Permuth [1,6,*,†]

[1] Department of Gastrointestinal (GI) Oncology, Moffitt Cancer Center, Tampa, FL
[2] Department of Bioinformatics and Biostatistics, Moffitt Cancer Center, Tampa, FL
[3] Division of Gastroenterology, Hepatology and Nutrition, The Ohio State University Comprehensive Cancer Center, The Ohio State University Wexner Medical Center, Columbus, OH
[4] Department of Pathology, Moffitt Cancer Center, Tampa, FL
[5] Department of Surgery, Division of Surgical Oncology, The Ohio State University Comprehensive Cancer Center, The Ohio State University Wexner Medical Center, Columbus, OH
[6] Department of Cancer Epidemiology, Moffitt Cancer Center, Tampa, FL
[7] UT Southwestern Medical Center, Dept. of Surgery, Dallas, TX
[8] Department of Machine Learning, Moffitt Cancer Center, Tampa, FL
[9] Department of Radiology, Moffitt Cancer Center, Tampa, FL
* Correspondence: Zobeida.Cruz-Monserrate@osumc.edu (Z.C.-M.); Jenny.Permuth@moffitt.org (J.B.P.)
† These authors have contributed equally to this work and share senior authorship.

Abstract: Introduction: Intraductal papillary mucinous neoplasms (IPMN) are commonly-detected pancreatic cysts that may transform into pancreatic ductal adenocarcinoma (PDAC). Predicting which IPMNs will progress to PDAC remains a clinical challenge. Moreover, identifying those clinically evident IPMNs for which a surveillance approach is best is a dire clinical need. Therefore, we aimed to identify molecular signatures that distinguished between PDAC with and without clinical evidence of a IPMN to identify novel molecular pathways related to IPMN-derived PDAC that could help guide biomarker development. **Methods:** Data from the Oncology Research Information Exchange Network *(ORIEN) multi-institute sequencing project was utilized to* analyze 66 PDAC cases from Moffitt Cancer Center and The Ohio State University Wexner Medical Center for which tumor whole transcriptome sequencing datasets were generated. Cases were classified based on whether a tumor had originated from an IPMN (n=16) or presumably through the pancreatic intraepithelial neoplasia (PanIN) pathway (n=50). We then performed differential expression and pathway analysis using Gene-Set Enrichment Analysis (GSEA) and Pathway Analysis with Down-weighted Genes (PADOG) algorithms. We also analyzed immune profiles using the Tumor-immune Microenvironment Deconvolution Web-portal for Bulk Transcriptomics (TIMEx). **Results:** Both GSEA and TIMEx indicate that PanIN-derived PDAC tumors enrich inflammatory pathways (complement, hedgehog signaling, coagulation, inflammatory response, apical surface, IL-2/STAT5, IL-6/STAT3, EMT, KRAS signaling, apical junction, IFN-gamma, allograft rejection) and are comparatively richer in almost all immune cell types than those from IPMN-derived PDAC upon deconvolution. IPMN-derived tumors were enriched in metabolic and energy generating pathways (oxidative phosphorylation, unfolded protein response, pancreas beta cells, adipogenesis, fatty acid metabolism, protein secretion). Further, the metabolic-linked gene signature enriched in the IPMN-derived samples is associated with a cluster of early stage and long survival (top 4th quartile) PDAC cases from the Cancer Genome Atlas (TCGA) expression database. Intriguingly, the gene most significantly upregulated for both the improved survival group and IPMN-derived PDAC was

 doi:10.20944/preprints202409.2045.v

gastrokine-2 (GKN2, padj<0.001), a gene upregulated in PDAC precursor lesions (compared to PDAC). **Conclusions:** Our data suggest that IPMN-derived and PanIN-derived PDACs differ in the expression of immune profiles and metabolic pathways. These initial findings support future studies to assess the accuracy of identifying PDACs arising from IPMN lesions based on either GKN2 expression or their metabolic or immune profiles and explore mechanisms to explain these findings.

Keywords: pancreatic ductal adenocarcinoma; IPMN; PanIN; transcriptomics

1. Introduction

Although great strides have been made in improving survival for many cancer types, the prognosis for pancreatic ductal adenocarcinoma (PDAC) remains grim, with a 5-year relative survival rate of only 13%[1–4]. Moreover, both incidence and mortality for PDAC are rising and this disease is projected to become the second leading cause of cancer-related mortality by 2030 [3–5]. One of the limitations to improving PDAC outcomes is the lack of effective early detection and prevention methods[6].

Advancing the field of PDAC early detection/prevention lies in studying individuals with diseases at a high-risk of developing PDAC, such as those with pre-malignant pancreatic cystic lesions known as intraductal papillary neoplasms (IPMNs) and mucinous cystic neoplasms (MCN), a sentiment in line with the National Cancer Institute's initiative called the 'Pre-Cancer Genome Atlas' which includes pancreatic cancers[6,7]. IPMNs are incidentally diagnosed in 13-45% of patients undergoing magnetic resonance imaging (MRI) and 2% of computed tomography (CT) scans[6,8–10]. Accurate diagnosis and risk-stratification of IPMNs has multiple direct and indirect clinical consequences such as appropriate application of curative intervention or alternatively attenuation of surgery-associated morbidities such as diabetes [11–15].

In addition to early detection, prevention of IPMN progression is an important avenue for clinical intervention. For example, since obesity and smoking are risk factors for IPMN progression to PDAC[16], dietary and exercise interventions may be a fruitful avenue for investigation[17,18].

Pursuing both early detection and prevention will in turn minimize the cost of healthcare and utilization of healthcare resources.

IPMN-derived PDAC seems to be a distinct clinical entity versus PDACs derived from the PanIN pathway as the prognosis of IPMN-derived PDAC is often better than sporadic PDACs (not derived from or occurring concomitantly with IPMNs)[19]. Furthermore, IPMN-derived PDACs may be distinct biologically. Both IPMN-derived and PanIN-derived PDAC contain similar mutation profiles, namely RAS and TP53 driver mutations, although a difference in the prevalence of activating KRAS (for PanIN-derived/conventional) versus GNAS (for IPMN-derived) has been reported in the literature [20–24]. Although both cystic and non-cystic lesions can lead to PDAC tumors, it is unclear whether the resulting tumors have differential expression patterns. Current literature has reported DNA mutational analysis and has not focused on differences in RNA expression between IPMN-derived and PanIN-derived lesions using a transcriptomic approach [20–25], though ITGA2 and SDC1 have been identified as potential prognostic genetic biomarkers for IPMN progression to PDAC in a recent study [26]. Thus, expression of unique molecules could be leveraged to develop imaging and/or activity probes that could be used to detect IPMN/pre-PDAC in a screening population[24,27]. Furthermore, differential expression analysis of IPMN- vs PanIN-derived lesions may identify novel pathways for IPMN risk stratification via biopsy or cyst fluid.

The objective of this study was to compare, for the first time, the transcriptome of PDAC tumors arising from PanIN and IPMN-derived etiologies to better characterize PDAC tumors based on their etiology. To assess possible differences between the PanIN and IPMN-derived samples, we used bioinformatic approaches to conduct pathway analyses and immune profiling.

 doi:10.20944/preprints202409.2045.v

2. Methods

Study population: The study population included male and female adults (≥ age 18) who consented to the Total Cancer Care (TCC) protocol[28] at Moffitt Cancer Center (Moffitt, Tampa, Florida) and The Ohio State University (OSU, Columbus, Ohio) who underwent surgical resection of a pancreatic tumor between 2005 and 2020 and were pathologically confirmed to have a diagnosis of PDAC or related histology using ICD-O-3 codes 82553, 84503, 81403, 84803, 85003 and 84532. Hematoxylin and eosin (H&E)-stained slides and electronic medical records from eligible patients were analyzed by a pathologist to determine IPMN involvement. If a pathology report noted that invasive cancer was present in association with and/or histologically contiguous with an IPMN, the PDAC was considered to be derived from the IPMN. If there was a distance from the focus of invasion and the IPMN according to the pathology report, or if no IPMN was noted, we considered a possibility of *de novo* origination and therefore were classified as PanIN-derived. Neuroendocrine tumors and metastases of pancreatic primary tumors were excluded.

Sample Handling and RNA Extraction: Surgically resected pancreas tumor tissue was retrieved from the Moffitt and OSU institutional biobanks for all eligible cases. Snap-frozen tissue aliquots were the specimen type of choice; if frozen tissue was unavailable, formalin-fixed paraffin embedded (FFPE) tissue blocks were used. After pathological review to confirm the diagnosis, tissue specimens underwent nucleic acid extraction and sequencing at Aster Insights/M2Gen/HudsonAlpha (Huntsville, AL). For RNA isolation from frozen tissue, the Qiagen RNAeasy plus mini kit was used, generating 216 bp average insert size. For FFPE tissue, the Covaris Ultrasonication FFPE DNA/RNA kit was utilized to extract both DNA and RNA, generating 165 bp average insert size. DV200 (fraction of RNA fragments longer than 200 base pairs) was >20 and total RNA was >20ng for all samples except for 1 (5.7 ng RNA) which was analyzed prior to updated standard operating procedures (SOPs) in 2016. RNA sequencing (RNAseq) was performed using the Illumina TruSeq RNA Exome with single library hybridization, cDNA synthesis, library preparation, sequencing (100 bp paired reads) to a coverage of 100M total reads / 50M paired reads. Detailed SOPs for FFPE sample collection are included in **Supplemental Methods**.

RNAseq Analysis: Fastq files were aligned to the human genome (hs37d5) using STAR (v2.5.3a)[29]. Gene-level quantitation was performed using HTSeq (v0.6.1), [30] and QC was performed on the counts files using standard RNAseq quality metrics (% aligned, % intronic) and visualizations.

Differential Expression Analysis: HTSeq count files were evaluated for standard metrics and visualizations such as counts per sample. Differential expression was assessed in samples using various contrasts using the "DESeq2" R package and a simple 2 group model[31].

Pathway Analysis and Immune Profiling: Fragments per kilobase million (FPKM) values were subjected to gene set enrichment analysis (GSEA) using GenePattern [32,33] and using either a publicly available universal gene set (hallmark pathways), or an immune deconvoluting gene signature for PDAC (the TIMEx tool) [34,35]. For further confirmation, FPKM values from the ORIEN/Avatar dataset were analyzed using the Reactome Web Tool [36–38].

Hierarchical Clustering: As a validation dataset, RNA expression data from pancreatic adenocarcinoma (PAAD) dataset from The Cancer Genome Atlas (TCGA) was downloaded from the Genome Data Commons (GDC portal) along with clinical characteristics of the participants including survival, stage and grade[39]. Hierarchical clustering was performed on scaled expression values using manually curated gene signatures and visualized using the "complexHeatmaps" R package[40].

Statistical Analysis: Statistical analyses were performed using R. Kaplan Meier survival analyses were conducted using the "survival" package. Chi square tests were performed on proportion type data. Principal Component Analysis (PCA) was performed using base R and plotted using "ggbiplot"[41].

 doi:10.20944/preprints202409.2045.v

3. Results

Characteristics of the Analytic Cohort: RNAseq data was generated on tumor tissue from 139 eligible cases with a total of 66/139 samples included in the analytic cohort; of the 73 samples that failed QC, most (n=60) were from FFPE while 13 were from flash frozen (FF) tissues. Thus, the final analytic dataset includes 66 cases (50 from Moffitt and 16 from OSU) (**Figure 1**). Patients were an average of 68.3 years of age at diagnosis/time of resection (range 48-87, **Table 1**). Most cohort participants were males (n=37, 56%) and most were non-Hispanic white (n=58, 87%, **Table 1**). The majority of cases (n=50/66) had conventional PanIN derived PDAC histology and 16 were IPMN-derived.

Clustering was not observed upon principal component analysis among all samples: To ensure that institution (Moffitt vs. OSU)- and sample-type batch effects were minimal, we performed principal component (PC) analyses to assess clustering. PC plots (PC1 vs PC2) colored based on tumor derivation reveals no clustering according to whether the sample was IPMN or PanIN-derived (**Figure S1**)**.** No appreciable batch effects were observed. PC plots colored based on cancer etiology, sample type, sex, race, and ethnicity also show no obvious clustering (**Figure S1**). Combined, PC1 and PC2 account for a total of 40% of the variance in the dataset.

Gene expression profiles differ between IPMN-derived versus PanIN-derived PDACs: A total of 215 genes were significantly deregulated (154 upregulated and 61 downregulated, adjusted p-value <0.05) when comparing expression levels of IPMN-derived versus non IPMN-derived PDAC tumors (**Table S1**). MUC2 was one of the most significantly upregulated (~4fold change (FC) gene in patients whose PDAC arose from a IPMN versus a PanIN (**Figure 2A**)**.** Other genes of interest upregulated in IPMN-derived tumors include gastrokinin 2 (GKN2, FC=7.5-) gastrokinin 1, (FC=7.6), pyruvate dehydrogenase kinase (PDK4, FC=2.4), carbamoyl phosphate synthase 1 (CPS1, FC=2.6) and serine peptidase inhibitor Kazal type 4 (SPINK4, FC=4.3). Alkaline phosphatase placental (ALPP, FC=-4.3) and the predicted protein coding gene C6orf15 (FC=-5.7) were some of the most significantly upregulated genes in PanIN-derived tumors.

Gene expression profiles differ between PDAC participants based on overall survival: As a secondary analysis, participants were dichotomized into short and long survival times based on a median split (**Figure 2B** and **Table S2**). A total of 657 genes were significantly deregulated (421 upregulated and 236 downregulated). GKN2 is highly upregulated when survival times are longer than the median (FC=7.8) as is its paralog GKN1 (FC=7.8). Cyclin-dependent kinase 5 receptor 1 (CDK5R1, FC=-1.6), uroplakin 2 (UPK2, FC=-3.2) and high mobility group AT-hook 2 (HMGA2) are upregulated in the short survival group. A total of 53 genes overlap (having the same direction of change (positive or negative) and having a padj<0.05) between the long vs short survival and IPMN vs non IPMN-derived PDAC comparison analyses. See **Table S3** for fold-change and p-values for this subset of genes.

Survival times differ between IPMN-derived and non IPMN-derived PDAC patients: A Kaplan Meier survival analysis indicated that IPMN-derived participants had significantly longer survival times than PanIN-derived (p=0.046, median is 36 months (Confidence interval (CI):25 lower confidence interval limit (LCL) – Upper confidence interval limit (UCL) could not be calculated) for IPMN-derived and 27 months (CI:19 LCL – 37 UCL) for PanIN-derived, (**Figure S2**). See **Table S4** for a chi square analysis of the proportion of each group which had short versus long survival times based on a median split (not significant) as was done to dichotomize groups for differential gene expression in **Figure 2B** (p-value was not significant at 0.3).

Gene Set Enrichment Analysis indicates differential regulation of cellular metabolism and immune/inflammatory signaling in IPMN- vs PanIN-derived PDAC: Normalized gene set enrichment scores (NES) from the Moffitt-OSU ORIEN/Avatar cohort from the publicly available "HALLMARK" gene set, or the immune deconvoluting gene set for PDAC (**Figure 3A-B**, **text file S1**) developed using single cell transcriptomic data from Tumor Immune Single Cell Hub (TISCH) [42] across 16 solid cancer types including PDAC (TIMEx) [34] were calculated using the GenePattern webtool and GSEA pipelines and filtered based on p-values (**Figure 3A-B**). As shown in **Figure 3A**, Hallmark gene sets enriched in PanIN-derived PDACs include allograft rejection and IFN-gamma

 doi:10.20944/preprints202409.2045.v

response as well as epithelial to mesenchymal transition and apical surface pathways. Gene sets enriched in IPMN-derived PDACs include oxidative phosphorylation, unfolded protein response and adipogenesis. Immune cell type gene signatures from TIMEx were almost exclusively enriched in PDACs not derived from cystic lesions and include monocytes/macrophages, neutrophils, mast cells, dendritic cells and CD8+ T cells (**Figure 3B**). A foam plot from the Reactome Webtool [36–38] demonstrates an enrichment of inflammation/immune-related genes/cell signatures in PanIN-derived PDAC, while IPMN-derived PDAC show enrichment for metabolic/energy-related and cell cycle genes (**Figure 3C**).

A validation cohort (TCGA) partially supports our findings for metabolism-linked pathways: To validate the findings using the TCGA PAAD cohort, a publicly available dataset which analyze RNAseq from pancreatic tumor tissue [43–45] in a supervised analysis with our curated gene sets. After filtering to adenocarcinoma samples, manually curated gene sets (IPMN-derived and PanIN-derived) were then used to cluster (hierarchical clustering) patient samples in the TCGA dataset. Custom gene profiles were assigned as follows: genes in Hallmark pathways from GSEA that were enriched (p-value of <0.05) in either IPMN-derived tumors (oxidative phosphorylation, unfolded protein response, pancreas beta cells, adipogenesis, fatty acid metabolism, protein secretion) or PanIN-derived tumors (complement, hedgehog signaling, coagulation, inflammatory response, apical surface, IL-2/STAT5, IL-6/STAT3, EMT, KRAS signaling, apical junction, IFN-gamma, allograft rejection) were considered if they appeared in more than 1 pathway (see **Table S5**).

As shown in **Figure 4**, we noted a small cluster of TCGA PAAD patients having upregulated tumor expression for genes in the IPMN-derived gene set and these patients are almost exclusively stage I, low grade (G1) and have long (based on a quartile split) disease-specific survival times. In addition, we noted a cluster of patients who demonstrate a moderate upregulation of IPMN-derived genes, and these patients show more mixed stage, grade and survival characteristics. There was no notable PanIN-derived (immune gene) clustering based on the gene profiles from pathways identified in the ORIEN/Avatar cohort.

4. Discussion

We demonstrate that significant expression-level and pathway-level differences occur between IPMN-derived versus PanIN-derived PDAC tumors. We highlight that IPMN-derived PDAC exhibits changes in genes related to metabolism while PanIN-derived PDAC exhibits changes in genes related to immunology. These molecular profiles have not been identified in other studies of IPMN-derived PDAC [46,47] though lipid profiles differ between cystic lesion types [48]. These findings may have implications for prognosis or treatment for these tumor types as both metabolic dysfunction and inflammation are targetable pathways [49–58].

When comparing IPMN-derived versus non IPMN-derived PDAC at the gene level, we observed significant dysregulation in many metabolism-related genes such as PDK4 (increased) and ALPP (decreased). For GSEA, normalized enrichment scores favored the IPMN-derived PDAC for energy-sensing and energy-production pathways such as oxidative phosphorylation, lipogenesis/fatty acid metabolism and mammalian target of rapamycin complex 1 (MTORC1, not shown as this pathway approached significance), consistent with existing data [50,59–61] and with a cluster of TCGA PAAD patients who were early stage and had long survival times. Conversely, we find that PanIN-derived PDAC was more likely to be enriched for immune/inflammatory pathways such as complement signaling, the inflammatory response and IFN-γ signaling. Others have demonstrated similar metabolic and/or inflammatory pathway enrichment when subtyping lung [62,63], breast [49] and GI [52,64,65] malignancies, though links to favorable versus unfavorable prognosis were mixed. The response to immunotherapy has so far been marginal for PDAC patients[66–68]. Nevertheless, our data and others' [48,63,64] suggest that dysregulation of genes in energy generating pathways may be a favorable prognostic indicator in some solid tumor types.

There is a complex interplay of pro- and anti-inflammatory cytokines/immune cells driving PDAC progression and drug resistance [69–75] and our findings demonstrate that, in comparison to IPMN-derived PDAC, PanIN-derived PDAC have a richer immune/inflammatory component. Taken

 doi:10.20944/preprints202409.2045.v

together, our GSEA data suggest that energy metabolism dysregulation may be useful to stratify patients for prognosis or therapeutic regimens[76]. These data provide rationale that exercise and dietary interventions may be especially beneficial for patients with IPMN-derived PDAC [17,77–79]. Others have reported racial disparities in response to immunotherapy (with African Americans having a better response than Whites [80–82]), and racial differences in energy metabolism are well documented [83–87]. Thus, our findings also underscore the dire need for minority inclusion in future PDAC clinical trials and research efforts [82].

Interestingly, when we stratified by survival (high/low based on a median split), genes associated with longer survival include both GKN1 and GKN2. GKN2 was also one of the most upregulated genes in the IPMN-derived cohort which is in line with literature suggesting that IPMN-derived PDAC patients have longer survival times, possibly due to earlier diagnosis [88,89]. Notably, GKN2 increases sensitivity/cell death in response to reactive oxygen species in gastric cancer cell lines [90]. Others have demonstrated that these genes have low expression in tumor tissues and inhibit PDAC progression [91–93]. Not unexpectedly, cyclin dependent kinase CDK5R1 [94] is upregulated in the low survival group, as well as uroplakin 2 (UPK2) [95–97] which is upregulated in some urothelial cancer types, and high mobility group A2 (HMGA2) [98–100]. HMGA2 is upregulated in multiple cancers including pancreatic cancer and is associated with multiple pro-tumor pathologies; however, its mechanism of action remains elusive.

While we used a unique cohort of samples from two institutions, there are some limitations to this study. First, we were only able to obtain expression profiles for a small cohort of moslty treatment-naïve PDAC patients (n=66), mainly due to issues with FFPE samples not passing QC for RNA sequencing. In addition, we were only able to confirm related cystic lesions in 16 of the cases, in line with published data on the incidence of IPMN-derived PDAC [12–15]. Furthermore, given that there were only 6 Hispanic or Black/African American participants (race of 2 participants was unavailable), race and ethnicity could not be statistically considered in this cohort and thus the implications for metabolic vs inflammatory signaling enrichment by race and ethnicity is unknown. While the distribution of cases were predominantly early -stage tumors this supports our goal to conduct investigations that would be informative for early detection of PDAC.

5. Conclusions

Enrichment in energy metabolism pathway genes was associated with IPMN-derived lesions and a better prognosis in our analytic cohort and in TCGA PAAD cohort. Taken together, our results suggest that early adoption of exercise and dietary interventions or metabolic inhibitors may be of clinical use for patients with early IPMN [77–79]. Furthermore, as GKN2 is highly upregulated in IPMN-derived PDAC patients in our cohort, these patients may respond favorably to platinum or other ROS-generating therapies [90]. Our study also suggests that inflammatory and immune pathways may be plausible targets for PanIN-associated PDAC patients. Further studies are needed to better define metabolic/inflammatory signatures for non IPMN- versus IPMN-derived PDAC patients and to better understand the role of gastrokinins in pancreatic tumors.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org.

Author Contributions: The authors confirm contribution to the paper as follows: study conception and design: Z.C-M, J.B.P, M.A.P; data collection: M.A.P, K.F-G.,J.K.T.; analysis and interpretation of results: Z.C-M, J.B.P., M.A.P., K.F-G.,M.G-D., A.M.; manuscript preparation: Z.C-M, J.B.P, M.A.P, K.F-G.; manuscript editing: Z.C-M, J.B.P, M.A.P, K.F-G, S.G.K, M.C.G-D., S.B., P.A.H, M.E.D., M.F.G., T.L.B., S.R.M, A.K.L., J.B.F., A.M., B.A.C., K.J., A.K., D.J., D-T.C., P.A.S., J.K.T. All authors reviewed the results and approved the final version of the manuscript.

Funding: This investigation was supported by funding awarded ORIEN New Oncology Visionary Award (NOVA) to J. P and Z. C-M, the National Cancer Institute of the National Institutes of Health (R37-CA229810 to JP and DJ), The Tissue Core, Collaborative Data Services Core, Molecular Genomics Core, and the Biostatistics and Bioinformatics Shared Resource, at the H. Lee Moffitt Cancer Center & Research Institute, an NCI designated Comprehensive Cancer Center (P30-CA076292) and The Ohio State University Comprehensive Cancer Center

 doi:10.20944/preprints202409.2045.v

and the National Institutes of Health (P30 CA016058). The content is solely the responsibility of the authors and does not necessarily represent the official views of the sponsors or the participating institution.

Acknowledgments: The authors wish to acknowledge all the patients who participated in the Orien/Avatar study. In addition, the authors wish to acknowledge the M2Gen teams who facilitated these findings and the Total Cancer Care arm of the Biospecimen Services Shared Resource at The Ohio State University Comprehensive Cancer Center, Columbus, OH for specimen provisioning.

Conflicts of Interest: S.R.M. is a consultant for C2/Pentax and Steris which are not related to the current work. All other 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.

References

1. Tonini, V. and M. Zanni, *Pancreatic cancer in 2021: What you need to know to win.* World J Gastroenterol, 2021. **27**(35): p. 5851-5889.
2. McGuigan, A., et al., *Pancreatic cancer: A review of clinical diagnosis, epidemiology, treatment and outcomes.* World J Gastroenterol, 2018. **24**(43): p. 4846-4861.
3. https://www.cancer.org/research/cancer-facts-statistics/all-cancer-facts-figures/2023-cancer-facts-figures.html.
4. Siegel, R.L., A.N. Giaquinto, and A. Jemal, *Cancer statistics, 2024.* CA Cancer J Clin, 2024. **74**(1): p. 12-49.
5. Rahib, L., et al., *Projecting cancer incidence and deaths to 2030: the unexpected burden of thyroid, liver, and pancreas cancers in the United States.* Cancer Res, 2014. **74**(11): p. 2913-21.
6. Campbell, J.D., et al., *The Case for a Pre-Cancer Genome Atlas (PCGA).* Cancer Prev Res (Phila), 2016. **9**(2): p. 119-24.
7. Srivastava, S., et al., *PreCancer Atlas: Present and Future.* Cancer Prev Res (Phila), 2023. **16**(7): p. 379-384.
8. Moris, M., et al., *Association Between Advances in High-Resolution Cross-Section Imaging Technologies and Increase in Prevalence of Pancreatic Cysts From 2005 to 2014.* Clin Gastroenterol Hepatol, 2016. **14**(4): p. 585-593 e3.
9. Laffan, T.A., et al., *Prevalence of unsuspected pancreatic cysts on MDCT.* AJR Am J Roentgenol, 2008. **191**(3): p. 802-7.
10. Pinho, D.F., N.M. Rofsky, and I. Pedrosa, *Incidental pancreatic cysts: role of magnetic resonance imaging.* Top Magn Reson Imaging, 2014. **23**(2): p. 117-28.
11. Hines, O.J. and H.A. Reber, *Pancreatic surgery.* Curr Opin Gastroenterol, 2008. **24**(5): p. 603-11.
12. Salvia, R., et al., *Branch-duct intraductal papillary mucinous neoplasms of the pancreas: to operate or not to operate?* Gut, 2007. **56**(8): p. 1086-90.
13. Maguchi, H., et al., *Natural history of branch duct intraductal papillary mucinous neoplasms of the pancreas: a multicenter study in Japan.* Pancreas, 2011. **40**(3): p. 364-70.
14. Rodriguez, J.R., et al., *Branch-duct intraductal papillary mucinous neoplasms: observations in 145 patients who underwent resection.* Gastroenterology, 2007. **133**(1): p. 72-9; quiz 309-10.
15. Yamada, S., et al., *Comparison of the international consensus guidelines for predicting malignancy in intraductal papillary mucinous neoplasms.* Surgery, 2016. **159**(3): p. 878-84.
16. Gentiluomo, M., et al., *Role of pancreatic ductal adenocarcinoma risk factors in intraductal papillary mucinous neoplasm progression.* Front Oncol, 2023. **13**: p. 1172606.
17. Pita-Grisanti, V., et al., *Physical Activity Decreases Inflammation and Delays Development of Obesity-Associated Pancreatic Ductal Adenocarcinoma.* Cancer Res, 2024.
18. Hildebrand, N.D., et al., *Supervised Home-Based Exercise Prehabilitation in Unfit Patients Scheduled for Pancreatic Surgery: Protocol for a Multicenter Feasibility Study.* JMIR Res Protoc, 2023. **12**: p. e46526.

 doi:10.20944/preprints202409.2045.v

19. Sahora, K., et al., *Branch duct intraductal papillary mucinous neoplasms: does cyst size change the tip of the scale? A critical analysis of the revised international consensus guidelines in a large single-institutional series.* Ann Surg, 2013. **258**(3): p. 466-75.
20. Turner, R.C., et al., *Molecular Analysis of Pancreatic Cyst Fluid for the Management of Intraductal Papillary Mucinous Neoplasms.* Diagnostics (Basel), 2022. **12**(11).
21. Hosoda, W., et al., *GNAS mutation is a frequent event in pancreatic intraductal papillary mucinous neoplasms and associated adenocarcinomas.* Virchows Arch, 2015. **466**(6): p. 665-74.
22. Tan, M.C., et al., *GNAS and KRAS Mutations Define Separate Progression Pathways in Intraductal Papillary Mucinous Neoplasm-Associated Carcinoma.* J Am Coll Surg, 2015. **220**(5): p. 845-854 e1.
23. Eshleman, J.R., et al., *KRAS and guanine nucleotide-binding protein mutations in pancreatic juice collected from the duodenum of patients at high risk for neoplasia undergoing endoscopic ultrasound.* Clin Gastroenterol Hepatol, 2015. **13**(5): p. 963-9 e4.
24. Distler, M., et al., *Precursor lesions for sporadic pancreatic cancer: PanIN, IPMN, and MCN.* Biomed Res Int, 2014. **2014**: p. 474905.
25. Huang, X., et al., *The molecular, immune features, and risk score construction of intraductal papillary mucinous neoplasm patients.* Front Mol Biosci, 2022. **9**: p. 887887.
26. Zhang, C.L., et al., *SDC1 and ITGA2 as novel prognostic biomarkers for PDAC related to IPMN.* Sci Rep, 2023. **13**(1): p. 18727.
27. Yoon, J.G., et al., *Pancreatic cystic neoplasms: a review of current recommendations for surveillance and management.* Abdom Radiol (NY), 2021. **46**(8): p. 3946-3962.
28. Fenstermacher, D.A., et al., *Implementing personalized medicine in a cancer center.* Cancer J, 2011. **17**(6): p. 528-36.
29. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.
30. Anders, S., P.T. Pyl, and W. Huber, *HTSeq--a Python framework to work with high-throughput sequencing data.* Bioinformatics, 2015. **31**(2): p. 166-9.
31. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.
32. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.
33. Mootha, V.K., et al., *PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes.* Nat Genet, 2003. **34**(3): p. 267-73.
34. Xie, M., et al., *TIMEx: tumor-immune microenvironment deconvolution web-portal for bulk transcriptomics using pan-cancer scRNA-seq signatures.* Bioinformatics, 2021.
35. Moffitt, R.A., et al., *Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma.* Nat Genet, 2015. **47**(10): p. 1168-78.
36. Gillespie, M., et al., *The reactome pathway knowledgebase 2022.* Nucleic Acids Res, 2022. **50**(D1): p. D687-D692.
37. Fabregat, A., et al., *Reactome diagram viewer: data structures and strategies to boost performance.* Bioinformatics, 2018. **34**(7): p. 1208-1214.
38. Griss, J., et al., *ReactomeGSA - Efficient Multi-Omics Comparative Pathway Analysis.* Mol Cell Proteomics, 2020. **19**(12): p. 2115-2125.
39. *Results based upon data generated by the TCGA Research Network: https://www.cancer.gov/tcga.*
40. Gu, Z., *Complex heatmap visualization.* Imeta, 2022. **1**(3): p. e43.
41. *https://github.com/friendly/ggbiplot.*

 doi:10.20944/preprints202409.2045.v

42. Sun, D., et al., *TISCH: a comprehensive web resource enabling interactive single-cell transcriptome visualization of tumor microenvironment.* Nucleic Acids Res, 2021. **49**(D1): p. D1420-D1430.
43. Cerami, E., et al., *The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.* Cancer Discov, 2012. **2**(5): p. 401-4.
44. Gao, J., et al., *Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.* Sci Signal, 2013. **6**(269): p. pl1.
45. de Bruijn, I., et al., *Analysis and Visualization of Longitudinal Genomic and Clinical Data from the AACR Project GENIE Biopharma Collaborative in cBioPortal.* Cancer Res, 2023. **83**(23): p. 3861-3867.
46. Sato, N., et al., *Gene expression profiling identifies genes associated with invasive intraductal papillary mucinous neoplasms of the pancreas.* Am J Pathol, 2004. **164**(3): p. 903-14.
47. Huang, B., et al., *Molecular characterization of organoids derived from pancreatic intraductal papillary mucinous neoplasms.* J Pathol, 2020. **252**(3): p. 252-262.
48. Gaiser, R.A., et al., *Integrated targeted metabolomic and lipidomic analysis: A novel approach to classifying early cystic precursors to invasive pancreatic cancer.* Sci Rep, 2019. **9**(1): p. 10208.
49. Tang, Y., et al., *Prognosis and Dissection of Immunosuppressive Microenvironment in Breast Cancer Based on Fatty Acid Metabolism-Related Signature.* Front Immunol, 2022. **13**: p. 843515.
50. Swierczynski, J., A. Hebanowska, and T. Sledzinski, *Role of abnormal lipid metabolism in development, progression, diagnosis and therapy of pancreatic cancer.* World J Gastroenterol, 2014. **20**(9): p. 2279-303.
51. Ose, J., et al., *Targeted Plasma Metabolic Profiles and Risk of Recurrence in Stage II and III Colorectal Cancer Patients: Results from an International Cohort Consortium.* Metabolites, 2021. **11**(3).
52. Gong, Y., et al., *Development of a prognostic metabolic signature in stomach adenocarcinoma.* Clin Transl Oncol, 2022. **24**(8): p. 1615-1630.
53. Chakraborty, P., et al., *Gene expression profiles of inflammatory breast cancer reveal high heterogeneity across the epithelial-hybrid-mesenchymal spectrum.* Transl Oncol, 2021. **14**(4): p. 101026.
54. El Rayes, T., et al., *Lung inflammation promotes metastasis through neutrophil protease-mediated degradation of Tsp-1.* Proc Natl Acad Sci U S A, 2015. **112**(52): p. 16000-5.
55. Wen, Y., et al., *Targeting Macrophage Migration Inhibitory Factor in Acute Pancreatitis and Pancreatic Cancer.* Front Pharmacol, 2021. **12**: p. 638950.
56. Geng, X., et al., *Cancer-Associated Fibroblast (CAF) Heterogeneity and Targeting Therapy of CAFs in Pancreatic Cancer.* Front Cell Dev Biol, 2021. **9**: p. 655152.
57. Pandey, V., et al., *CXCL10/CXCR3 signaling contributes to an inflammatory microenvironment and its blockade enhances progression of murine pancreatic precancerous lesions.* Elife, 2021. **10**.
58. Pointer, D.T., Jr., et al., *Neutrophil to lymphocyte ratio, not platelet to lymphocyte or lymphocyte to monocyte ratio, is predictive of patient survival after resection of early-stage pancreatic ductal adenocarcinoma.* BMC Cancer, 2020. **20**(1): p. 750.
59. Huang, X., et al., *LINC00842 inactivates transcription co-regulator PGC-1alpha to promote pancreatic cancer malignancy through metabolic remodelling.* Nat Commun, 2021. **12**(1): p. 3830.
60. Nimmakayala, R.K., et al., *Metabolic programming of distinct cancer stem cells promotes metastasis of pancreatic ductal adenocarcinoma.* Oncogene, 2021. **40**(1): p. 215-231.
61. Nimmakayala, R.K., et al., *PGC1alpha-Mediated Metabolic Reprogramming Drives the Stemness of Pancreatic Precursor Lesions.* Clin Cancer Res, 2021. **27**(19): p. 5415-5429.
62. Huang, J., et al., *Tumor microenvironment characterization identifies two lung adenocarcinoma subtypes with specific immune and metabolic state.* Cancer Sci, 2020. **111**(6): p. 1876-1886.

 doi:10.20944/preprints202409.2045.v

63. Stewart, P.A., et al., *Proteogenomic landscape of squamous cell lung cancer.* Nat Commun, 2019. **10**(1): p. 3578.
64. Yang, C., et al., *Metabolism-associated molecular classification of hepatocellular carcinoma.* Mol Oncol, 2020. **14**(4): p. 896-913.
65. Pan, Y., et al., *Development and validation of robust metabolism-related gene signature in the prognostic prediction of hepatocellular carcinoma.* J Cell Mol Med, 2023. **27**(7): p. 1006-1020.
66. Pihlak, R., et al., *Advances in Molecular Profiling and Categorisation of Pancreatic Adenocarcinoma and the Implications for Therapy.* Cancers (Basel), 2018. **10**(1).
67. Smith, H., et al., *Manipulating regulatory T cells: is it the key to unlocking effective immunotherapy for pancreatic ductal adenocarcinoma?* Front Immunol, 2024. **15**: p. 1406250.
68. Sarfraz, Z., et al., *The Current Landscape of Clinical Trials for Immunotherapy in Pancreatic Cancer: A State-of-the-Art Review.* J Gastrointest Cancer, 2024. **55**(3): p. 1026-1057.
69. Helm, O., et al., *Tumor-associated macrophages exhibit pro- and anti-inflammatory properties by which they impact on pancreatic tumorigenesis.* Int J Cancer, 2014. **135**(4): p. 843-61.
70. Karakhanova, S., et al., *Characterization of myeloid leukocytes and soluble mediators in pancreatic cancer: importance of myeloid-derived suppressor cells.* Oncoimmunology, 2015. **4**(4): p. e998519.
71. Feng, R., et al., *Nab-paclitaxel interrupts cancer-stromal interaction through C-X-C motif chemokine 10-mediated interleukin-6 downregulation in vitro.* Cancer Sci, 2018. **109**(8): p. 2509-2519.
72. McAllister, F., et al., *Oncogenic Kras activates a hematopoietic-to-epithelial IL-17 signaling axis in preinvasive pancreatic neoplasia.* Cancer Cell, 2014. **25**(5): p. 621-37.
73. Zhao, X., et al., *Inhibiting tumor necrosis factor-alpha diminishes desmoplasia and inflammation to overcome chemoresistance in pancreatic ductal adenocarcinoma.* Oncotarget, 2016. **7**(49): p. 81110-81122.
74. Senturk, Z.N., et al., *Pancreatic cancer: Emerging field of regulatory B-cell-targeted immunotherapies.* Front Immunol, 2023. **14**: p. 1152551.
75. Halbrook, C.J., et al., *Macrophage-Released Pyrimidines Inhibit Gemcitabine Therapy in Pancreatic Cancer.* Cell Metab, 2019. **29**(6): p. 1390-1399 e6.
76. Shi, J., et al., *Cyst fluid metabolites distinguish malignant from benign pancreatic cysts.* Neoplasia, 2021. **23**(11): p. 1078-1088.
77. Fhu, C.W. and A. Ali, *Fatty Acid Synthase: An Emerging Target in Cancer.* Molecules, 2020. **25**(17).
78. Zhang, S., Y. Wei, and C. Wang, *Impacts of an Exercise Intervention on the Health of Pancreatic Beta-Cells: A Review.* Int J Environ Res Public Health, 2022. **19**(12).
79. Agarwal, S., et al., *AMP-activated Protein Kinase (AMPK) Control of mTORC1 Is p53- and TSC2-independent in Pemetrexed-treated Carcinoma Cells.* J Biol Chem, 2015. **290**(46): p. 27473-86.
80. Gupta, A., et al., *Racial Differences in Survival Among Advanced-stage Non-small-Cell Lung Cancer Patients Who Received Immunotherapy: An Analysis of the US National Cancer Database (NCDB).* J Immunother, 2022. **45**(2): p. 132-137.
81. Olateju, O.A., et al., *Investigation of racial differences in survival from non-small cell lung cancer with immunotherapy use: A Texas study.* Front Oncol, 2022. **12**: p. 1092355.
82. Nazha, B., et al., *Enrollment of Racial Minorities in Clinical Trials: Old Problem Assumes New Urgency in the Age of Immunotherapy.* Am Soc Clin Oncol Educ Book, 2019. **39**: p. 3-10.
83. Rahman, M., et al., *Racial differences in body fat distribution among reproductive-aged women.* Metabolism, 2009. **58**(9): p. 1329-37.
84. Racette, S.B., et al., *Racial differences in lipid metabolism in women with abdominal obesity.* Am J Physiol Regul Integr Comp Physiol, 2000. **279**(3): p. R944-50.

 doi:10.20944/preprints202409.2045.v

85. Hunter, G.R., et al., *Racial differences in energy expenditure and aerobic fitness in premenopausal women.* Am J Clin Nutr, 2000. **71**(2): p. 500-6.
86. Kimm, S.Y., et al., *Racial differences in the relation between uncoupling protein genes and resting energy expenditure.* Am J Clin Nutr, 2002. **75**(4): p. 714-9.
87. Berman, D.M., et al., *Racial disparities in metabolism, central obesity, and sex hormone-binding globulin in postmenopausal women.* J Clin Endocrinol Metab, 2001. **86**(1): p. 97-103.
88. de la Fuente, J., et al., *Long-Term Outcomes and Risk of Pancreatic Cancer in Intraductal Papillary Mucinous Neoplasms.* JAMA Netw Open, 2023. **6**(10): p. e2337799.
89. Koh, Y.X., et al., *Systematic review and meta-analysis comparing the surgical outcomes of invasive intraductal papillary mucinous neoplasms and conventional pancreatic ductal adenocarcinoma.* Ann Surg Oncol, 2014. **21**(8): p. 2782-800.
90. Zhang, Z., et al., *GKN2 promotes oxidative stress-induced gastric cancer cell apoptosis via the Hsc70 pathway.* J Exp Clin Cancer Res, 2019. **38**(1): p. 338.
91. Steiner, S., et al., *De novo expression of gastrokines in pancreatic precursor lesions impede the development of pancreatic cancer*. Oncogene, 2022. **41**(10): p. 1507-1517.
92. Du, J.J., et al., *[Down-regulated full-length novel gene GDDR and its effect on gastric cancer].* Zhonghua Yi Xue Za Zhi, 2003. **83**(13): p. 1166-8.
93. Yang, X., et al., *Low-level gastrokine 2 promoted progress of NSCLC and as a potential biomarker.* J Clin Lab Anal, 2022. **36**(2): p. e24213.
94. Dastjerdi, S., et al., *Elevated CDK5R1 expression associated with poor prognosis, proliferation, and drug resistance in colorectal and breast malignancies: CDK5R1 as an oncogene in cancers.* Chem Biol Interact, 2022. **368**: p. 110190.
95. Li, W., et al., *Uroplakin II is a more sensitive immunohistochemical marker than uroplakin III in urothelial carcinoma and its variants.* Am J Clin Pathol, 2014. **142**(6): p. 864-71.
96. Tajima, S. and K. Koda, *Uroplakin II Expression in Breast Carcinomas Showing Apocrine Differentiation: Putting Some Emphasis on Invasive Pleomorphic Lobular Carcinoma as a Potential Mimic of Urothelial Carcinoma at Metastatic Sites.* Dis Markers, 2016. **2016**: p. 2940496.
97. Wu, X., et al., *Uroplakin II as a promising marker for molecular diagnosis of nodal metastases from bladder cancer: comparison with cytokeratin 20.* J Urol, 2005. **174**(6): p. 2138-42, discussion 2142-3.
98. Fedele, M., et al., *Critical role of the HMGA2 gene in pituitary adenomas.* Cell Cycle, 2006. **5**(18): p. 2045-8.
99. Watanabe, S., et al., *HMGA2 maintains oncogenic RAS-induced epithelial-mesenchymal transition in human pancreatic cancer cells.* Am J Pathol, 2009. **174**(3): p. 854-68.
100. Wang, X., et al., *Overexpression of HMGA2 promotes metastasis and impacts survival of colorectal cancers.* Clin Cancer Res, 2011. **17**(8): p. 2570-80.