Article type: Original research

Manuscript title:

Glycation Stress–Driven Transcriptomic Signature Predicts Survival Benefit from Adjuvant Gemcitabine in Resectable Pancreatic Cancer.

Short title

Methylglyoxal-stress classifier for PDAC sensitivity to gemcitabine.

List of authors

Oier Azurmendi Senar[1], Kosta Stosic[1], Christelle Bouchart[1,2], Julie Navez[1,3], Jawad Tarfouss[1], Rebekah Crake[4], Laurine Verset[5], Pieter Demetter[1], Marjorie Mauduit[6], Thierry Conroy[7], Jérôme Cros[8,9], Remy Nicolle[9] , Manuel Saiselet[10], Joel Rodrigues Vitoria[10], Vincent Detours[10], Akeila Bellahcène[4 §], Tatjana Arsenijevic[1,11§] and Jean-Luc Van-Laethem[1,11§]; the PRODIGE-24/CCTG PA6 consortium; the MOSAPAC consortium.

[§]Authors contributed equally

[1]Laboratory of Experimental Gastroenterology, Université Libre de Bruxelles, Route de Lennik 808, 1070 Brussels, Belgium.

[2]Department of Radiation Oncology, HUB Institute Jules Bordet, Université Libre de Bruxelles, Rue Meylenmeersch 90, 1070 Brussels, Belgium.

[3]Department of Hepato-biliary-pancreatic surgery, Erasme University Hospital, Université Libre de Bruxelles, Brussels, Belgium.

[4] Metastasis Research Laboratory, GIGA-Cancer, GIGA Institute, University of Liège, Belgium.

[5] Department of Pathology, HUB Bordet Erasme Hospital, Université Libre de Bruxelles, Rue 21 Meylenmeersch 90, 1070 Brussels, Belgium.

[6] R&D Unicancer, Paris, France.

[7] Medical Oncology Department, Institut de Cancérologie de Lorraine, 54519 Vandoeuvre-lès-Nancy, France and Université de Lorraine, Inserm INSPIIRE, 54000, Nancy, France.

[8] Université Paris Cité, Centre de Recherche sur l'Inflammation (CRI), INSERM, U1149, CNRS, ERL 8252, F-75018 Paris, France.

[9] Université Paris Cité, Department of Pathology, FHU MOSAIC, Beaujon/Bichat University Hospital (AP-HP), Paris, France.

[10] Institute of Interdisciplinary Research (IRIBHM) Jacques E. Dumont, Université Libre de Bruxelles, Brussels, Belgium.

[11] Department of Gastroenterology, Hepatology and Digestive Oncology, HUB Bordet Erasme Hospital, Université Libre de Bruxelles, Route de Lennik 808, 1070 Brussels, Belgium.

Acknowledgments

We thank Unicancer and MOSAPAC consortia for providing PRODIGE-24/CCTG PA6 clinical data.

Abstract

BACKGROUND & AIMS

Pancreatic ductal adenocarcinoma (PDAC) is a highly lethal cancer with limited response to systemic therapy. Gemcitabine (GEM) benefits only a subset of patients. Methylglyoxal (MG), a glycolysis byproduct, has been linked to tumor behavior and therapy response in PDAC, suggesting potential as a stratification marker.

METHODS

We developed a metabolically informed gene signature (MG-GEM) integrating MG-related glycolytic stress with clinical outcomes. Using the Puleo cohort (n=309), differential expression analysis between tumors with high and low MG stress identified 365 genes. LASSO Cox regression selected 16 prognostic genes, combined into a weighted risk score for patient stratification. MG GEM was validated in internal and external cohorts, including PRODIGE 24/CCTG PA6. Molecular, transcriptomic, and immune features were compared between high and low MG GEM groups. Finally, predictive performance was evaluated against the GemPred signature.

RESULTS

MG GEM divided PDAC patients into distinct risk groups with marked differences in overall and disease-free survival among GEM treated patients (OS 11.7 vs 27.2 months; DFS 7.6 vs 17.8 months, both $p<0.0001$). High MG-GEM tumors showed enrichment for KRAS G12D and SMAD4 mutations, basal and activated stroma subtypes, glycolytic metabolism, and reduced immune infiltration. Low MG-GEM tumors showed KRAS G12V, classical and immune subtypes, cholesterogenic metabolism, and adaptive immune favourable signatures. MG-GEM independently predicted GEM-specific clinical outcomes, irrespective of GemPred signature, and significantly enhanced patient stratification when combined with it. Within the PRODIGE-24/TGCC PA6 cohort, MG-GEM exhibited prognostic relevance and selectively identified patients who derived a survival benefit from adjuvant GEM, but not from FOLFIRINOX.

CONCLUSIONS

The 16 gene MG GEM signature predicts prognosis in resected PDAC, reflects glycolytic stress driven chemoresistance, surpassing conventional molecular classifications. As a metabolically informed signature, MG-GEM holds promise for guiding chemotherapy selection and informing KRAS-targeted combination strategies, meriting further prospective clinical validation.

Introduction

Pancreatic ductal adenocarcinoma (PDAC), projected to become the second leading cause of cancer-related death by 2030 [1], is highly aggressive, with a 5-year overall survival (OS) rate below 12% [2]. Progress in developing more effective chemotherapy (CT) regimens has been limited [3], [4], [5] and clinical trials testing immunotherapy or targeted agents have, so far, failed to change the current clinical practice [6], [7] Currently, two CT regimens form the backbone of PDAC treatment: full-dose or modified FOLFIRINOX (folinic acid, 5-fluorouracil, irinotecan, and oxaliplatin; mFFX) and gemcitabine combined with nab-paclitaxel (GEM/nab-P) [8]. Their efficacy is broadly comparable, and choice between them is primarily guided by clinical criteria.

Gemcitabine (GEM) remains a cornerstone of systemic therapy for PDAC, yet only a subset of patients derives meaningful benefit. To address this challenge, transcriptomic classifiers have been developed to stratify patients based on their likelihood of responding to GEM. One such signature, GemPred [9], derived from preclinical models including cell lines and organoids, identifies approximately 26% of PDAC patients as predicted responders to GEM-based regimens. Additional transcriptomic tools have also been proposed to guide therapeutic decision-making. For example, GemCore [10], an integrative signature combining four independent GEM-response classifiers, identifies ~45% of patients as GEM-sensitive. Similarly, Pancreas-View [11], which integrates transcriptomic data from preclinical models with machine learning to provide PDAC molecular subtyping and therapy-response predictions, identifies ~37% of patients as GEM-sensitive. Together, these efforts underscore the ongoing commitment to optimizing precision treatment for PDAC.

In recent years, accumulating evidence has underscored the critical role of reprogrammed metabolism in shaping cancer treatment response and prognosis. In pancreatic cancer cells, genetic mutations are notably implicated in driving distinct metabolic phenotypes, including the Warburg effect [12]. Through this altered metabolic state, cancer cells increase glucose uptake and preferentially ferment glucose to lactate, even in the presence of oxygen [13]. Enhanced glycolytic activity inevitably results in the production of methylglyoxal (MG), a highly reactive aldehyde. Its accumulation and the formation of associated protein adduct, observed at elevated levels in breast, colon, and pancreatic cancers [14], underscore its significant role in tumor progression and chemoresistance.

In PDAC, previous studies have also linked glycolytic metabolism to chemoresistance and poor clinical outcomes [15], [16]. Aberrant KRAS activation, a hallmark of PDAC, drives enhanced glucose uptake and glycolysis through the upregulation of GLUT1 and key metabolic enzymes [12], [17]. A recent mechanistic study demonstrated that

increased glycolytic flux in PDAC models supports de novo pyrimidine biosynthesis, thereby reducing GEM efficacy through a competitive metabolic mechanism [18].Interestingly, GEM-resistant cancer cells display elevated levels of glycolytic intermediates upstream of dihydroxyacetone phosphate and glyceraldehyde-3-phosphate [18], both of which serve as precursors for the spontaneous formation of MG [19]. Moreover, we have shown in PDAC cell models that GEM treatment itself induces accumulation of MG-adducts, and resistant cells adapt to MG stress by upregulating GLO1, the key MG-detoxifying enzyme [20]. Together, these findings highlight the biological relevance of MG stress in GEM resistance and support a metabolism-focused scoring strategy centred on MG, which we found to correlate with poorer overall survival in GEM-treated PDAC patients [20]. To operationalize this, MG stress levels in chemotherapy-naïve resected PDAC tumors were assessed using two complementary indicators: expression of GLO1 and of a panel of ten glycolysis-related genes. This approach enabled the classification of PDAC tumors into three MG stress categories: high MG stress, low MG stress, and Intermediate. Although the MG stress score was prognostic, its utility was limited to a small subset of the cohort we explored, as majority of tumors fell into the intermediate category [20].

This study aimed to refine the MG-stress score into a prognostically optimized gene signature that integrates MG stress–associated biology with prediction of clinical outcome in GEM receiving resectable pancreatic cancer. To achieve this, we applied LASSO Cox regression to 365 differentially expressed genes between high- and low-MG stress tumors, identifying a subset of genes most strongly associated with prognosis. These genes were combined into a weighted risk score, forming the 16-gene MG-GEM signature, which enables patient stratification based on predicted clinical outcomes.

We further explored the clinical and therapeutic relevance of the MG-GEM signature in both internal and external cohorts, demonstrating its predictive value as a tool for selecting PDAC patients likely to benefit from adjuvant GEM treatment.

Materials and Methods

Patient cohorts and study design

This study analyzed three cohorts of patients with pancreatic ductal adenocarcinoma (PDAC) using the following criteria: inclusion of patients ⩾18 years old with complete clinicopathological data and no preoperative metastatic disease; exclusion of non-ductal histology and death within 30 days of surgery.

1. Puleo et al. cohort (n = 309, 1999–2010): Treatment-naïve, surgically resected PDAC cases with comprehensive clinical annotations and microarray-based

transcriptomic data. The 57.28% (n=177) of patients of the cohort received adjuvant GEM, 31.71% (n=98) of the patients did not receive any adjuvant treatment (No AT) and the 7.76% (n=24) received other adjuvant strategies (Other AT). A subset of 89 patients (GEM: 64.05% (n=57); No AT: 28.09% (n=25); Other AT: 4.5% (n=4)) served as the discovery set for the derivation of the MG-GEM signature [21], and 220 patients (GEM: 54.55% (n=120); No AT: 33.18% (n=73); Other AT: 9.09% (n=20)) were used for the signature validation (Figure 1-A).

2. Erasme University Hospital cohort (n = 60, 2011–2020): Retrospective cohort of treatment-naïve, surgically resected PDAC cases with available FFPE tumor specimens for transcriptome profiling [22]. Most patients received adjuvant chemotherapy, and detailed follow-up allowed assessment of the prognostic and predictive value of the MG-GEM signature.

3. PRODIGE24/CCTG PA6 phase III trial cohort (n = 350; EudraCT: 2011-002026-52): Randomized trial comparing modified FOLFIRINOX versus GEM as adjuvant therapy [23], providing an external validation dataset to evaluate the predictive performance of MG-GEM for GEM survival relative to FOLFIRINOX.

Figure 1-A shows a CONSORT-style flow diagram detailing patient cohorts used in this study for MG-GEM model generation and validation.

Data acquisition and processing

The discovery cohort consisted of 309 samples from the Puleo et al. study [21], profiled on Affymetrix HGU219 RNA microarrays, with probe-level intensities processed using Robust Multi-array Average (RMA) normalization. For the PRODIGE-24/CCTG PA6 dataset [3], raw gene counts were normalized via the Upper Quartile method and subsequently $\log_2$-transformed prior to downstream analyses.

Transcriptomic profile process and analysis

Patients from the Puleo cohort (n=309) [21] were classified into high (n=44) and low (n=45) MG stress tumors using the method of Crake et al. [20] (Figure 1-B). This classification combines the expression of 10 glycolytic genes with that of GLO1, the main MG-detoxifying enzyme. Differential gene expression (DGE) analysis was performed between patients with different MG-stress status using the R-packages edgeR v3.40.2 and limma v3.54.2 packages [24], [25]. Genes were considered differentially expressed when log two-fold change was at least ±1 and p-adjusted value < 0.05. The Pancreatic Adenocarcinoma Molecular Gradient (PAMG) classifier was applied to determine the aggressiveness of the samples [26].

Construction of the MG-GEM signature

Samples used for DGE analysis were randomly separated into a training set (75%) and a testing set (25%). Least Absolute Shrinkage and Selection Operator (LASSO) penalized Cox regression was conducted using the glmnet v4.1-8 package [27]. The RMA-normalized expression values of the differentially expressed genes (DEG) of the training cohort were used as input to fit the LASSO Cox regression model with the OS as the response variable. The survival v3.5-3 package was used to create the survival object, which acted as the response variable. Parameter alpha was indicated as 1 to ensure the application of the lasso penalty. The goal of this model was to extract the genes with a prognostic impact that were differentially expressed across MG-stress groups by constructing a penalty function to the residual sum of squares (RSS), which was then multiplied by the regularization parameter lambda (λ). The level of the regularization of the model is controlled by the hyperparameter lambda (λ) value that balances the bias and variance in the resulting coefficients.

Henceforth, coefficient β values were obtained to identify the optimal genes using the 10-fold cross-validation. Based on the risk score of each sample, the cohort was divided into two groups (low-risk with 0–50% vs high-risk with 50–100%). The diagnostic value of these genes was evaluated using receiver operating characteristic (ROC) curves.

scRNA-seq processing pipeline

Primary tissue datasets (GSE154778, GSE155698, PRJCA001063) were scaled and normalized with SCTransform (Seurat v5.0.1) [28]. The 3,000 most variable features were used for PCA, and the top 30 components for SNN graph construction, clustering, and UMAP embedding. Cohorts were integrated with Harmony (Supplementary Figure 4) [29]. Cluster marker genes were identified with FindAllMarkers and annotated based on the literature.

Sample processing and transcriptomic profile of ST samples

ST profiling was performed as described by Tourneur et al. [30]. Frozen tumor blocks were sectioned (10 µm) and processed with the spRNA-seq protocol [31]. Sections were mounted on Codelink Activated slides, fixed, H&E-stained, scanned (40×), and permeabilized. After cDNA synthesis, tissue removal, probe release, and library preparation, samples were quality-checked and sequenced on an Illumina NovaSeq 6000. Bright-field H&E images were acquired with a Nanozoomer-SQ scanner (Hamamatsu C13140-01).

ST-seq processing pipeline

Nanozoomer-SQ images were processed with ndpi2tiff v1.8, tifffastcrop v1.3.10 and vips v8.4.5 to rotate, convert to pyramidal TIFF, and generate JPEGs. ST Spot Detector [32] was used for image alignment and spot detection. A custom affine transform, defined by four reference points on the full-size H&E tissue, interpolated spot coordinates to full-size XY positions. Raw sequencing data was processed with the Salmén et al. pipeline v1.7.6 [33] (alignment, annotation, demultiplexing, filtering) using GRCh38 and Gencode v28. Outputs (UMI counts, corrected spot coordinates, H&E images) were integrated into a Seurat spatial object. Data were normalized with SCTransform (Seurat v5.0.1) and processed following Seurat guidelines. Samples were merged, reduced by PCA, clustered, and annotated using FindAllMarkers. Gene signature scores were calculated with AddModuleScore.

Functional analysis

Gene Set Enrichment Analysis (GSEA) was performed on a pre-ranked list of genes using the fgsea R package v1.24.0. The enrichment of gene sets with a padj< 0.05 were considered as significant. Gene signatures of PDAC and cancer-associated fibroblasts (CAF) subtypes were collected from the CancerRNASig package. Molecular Signature Database (MSigDB), Ontology and Canonical pathways gene sets were obtained by the msigdb package v1.6.0. Finally, immune cell fractions were estimated by the MCP counter algorithm and statistical analysis between treatments of the immune populations was obtained by the package ggpubr v0.6.0.

Statistical analysis

Patient characteristics were summarized using descriptive statistics (counts, percentages). Group comparisons were performed using Fisher's exact or Pearson's Chi-squared test, depending on expected counts, with gtsummary v1.7.2.

Survival analyses were performed using the survival package (v3.5-3) and survminer package (v0.4.9). Kaplan–Meier curves were compared using log-rank tests ($p < 0.05$). Multivariate Cox proportional hazards models were used to estimate hazard ratios (95% CI). Overall survival (OS) was defined as the time from diagnosis to death, and disease-free survival (DFS) as the time from diagnosis to recurrence after surgery.Transcriptomic data were analyzed using the Wilcoxon test in R (v4.2.3) and RStudio (v2023.3.0.386), with statistical significance set at $p < 0.05$. Group comparisons for PAMG and gene expression were performed using the ggpubr package (v0.6.0). [34], [35].

Results

Identification of a 16-gene signature and its correlation with clinicopathological findings

Crake et al. [20] MG stress classification was applied to the Puleo cohort [21] (n=309) allowing the classification of PDAC tumors with MG high (n=45) or MG low (n=44) status, as detailed in the materials and methods section. Differential gene expression (DGE) analysis identified 365 differentially expressed genes between the two groups (Figure 1). Impact prognostic genes were then selected using a LASSO Cox regression model based on overall survival (OS) (Supplementary Figure 1-A). The optimal regularization parameter (λ) was chosen by ten-fold cross-validation at the λ minimum error (Supplementary Figure 1-B). Sixteen genes with non-zero coefficients were retained (*FAM162A, SEC61G, NT5E, TREM1, SLC3A1, CALB2, TFF1, SLC11A1, PDIA2, PADI1, ADH1C, FXYD3, SCEL, ECT2, TIPARP, IGHG1*) (Supplementary Table 1).

A risk score was calculated from the expression and coefficients of these genes. Based on the median score, patients were divided into MG-GEM high (n=43) and low (n=42) groups (Figure 2-A). The model showed high predictive accuracy (AUC: 0.854 at 2 years, 0.778 at 4 years, 0.753 at 6 years) (Supplementary Figure 1-C).

Univariate analysis of CP findings revealed that MG-GEM status was significantly associated with perineural invasion (high: 85%, low: 66%, p=0.040) (Table 1). In the discovery cohort, high MG-GEM patients had significantly shorter OS (*11.7 vs. 27.2 months, p<0.0001*) and DFS (*7.6 vs. 17.8 months, p<0.0001*) compared with low MG-GEM patients (Figure 2-B,2-C).

To test whether MG-GEM was an independent prognostic factor, a multivariate Cox model was constructed including histological grade, perineural invasion, and Puleo subtype. MG-GEM remained independently associated with OS (*HR 3.5, 95% CI 1.7–7.3, p<0.001*) (Figure 2-D).

Molecular and Transcriptomic Profiling Reveal Distinct Tumor Features Associated with MG-GEM Stratification

To define the molecular landscape of MG-GEM groups, we examined key mutations and transcriptomic subtypes (Supplementary Table 2). *KRAS G12D* was more frequent in high MG-GEM tumors (59% vs. 33%, p=0.043), while *G12V* predominated in low MG-GEM tumors (53% vs. 22% p=0.043) *SMAD4* nonsense mutations were enriched in high MG-GEM (25% vs. 7.9%, p=0.049).

Transcriptomic classifiers consistently linked high MG-GEM tumors to poor-prognosis subtypes. In the Puleo scheme, they were mainly Pure Basal-Like (20% vs 2.7%) and Stroma Activated (38% vs 14%), while low MG-GEM tumors were often Immune Classical (2.5% vs 32%). Similarly, Bailey classification identified most high MG-GEM tumors as squamous (67% vs 9.5%), whereas low MG-GEM tumors were largely progenitor (12% vs 48%) or immunogenic (2.3% vs 14%).

Moffitt stromal profiling showed more activated stroma in high MG-GEM (67% vs. 43%) and normal stroma in low (2.3% vs 19%) (p=0.014). PurIST classification further distinguished the groups: basal tumors were increased in high MG-GEM (33% vs. 8.1%), while classical tumors predominated in low MG-GEM (68% vs. 92%), (p=0.008). GSEA confirmed enrichment of “Basal-like” and “Stroma Active” in high MG-GEM, and “Classical,” “Immune,” and “Stroma Inactive” in low MG-GEM (Figure 3-A).

We also explored the utility of MG-GEM in relation to tumor aggressiveness using the Pancreatic Adenocarcinoma Molecular Gradient (PAMG) score [26]. Consistently, PAMG scores were higher in low MG-GEM tumors (p=2.4e-06), reflecting a less aggressive phenotype (Supplementary Figure 1-D). High MG-GEM tumors also showed enrichment of poor-outcome Hwang et al. [36] single-cell signatures (“Basaloid,” “Mesenchymal,” “Squamoid”), while low MG-GEM tumors were enriched in favorable ones (“Neuroendocrine-like,” “Classical-like”) (Figure 3-B).

Next, we further localized MG-GEM gene expression to tumor epithelial cells. In 7 spatial transcriptomics (ST) samples from Erasme repository, MG-GEM genes were highly expressed in basal, acinar, and B cell–associated clusters (Supplementary Figure 2-A) (Supplementary Figure 3). Integration of three scRNA-seq cohorts confirmed expression restricted to tumor epithelial cells (Supplementary Figure 2-B) (Supplementary Figure 4).

MG GEM is correlated with immune and metabolic contexture in High MG GEM tumors

MCP-counter deconvolution showed a higher abundance of fibroblasts in high MG-GEM tumors (Supplementary Figure 5-A). CAF subtype analysis revealed enrichment of Hwang et al. [41] Adhesive and Myofibroblastic CAFs (myCAFs) signatures in high MG-GEM, while low MG-GEM tumors were enriched in “Neurotropic” CAFs (Figure 3-C). Using Kieffer et al. [37] signatures, GSEA confirmed predominance of myCAF programs in high MG-GEM (“TGFβ signaling,” “ECM proteins”), whereas low MG-GEM samples were enriched in normal fibroblast and inflammatory CAF (iCAF) signatures (Figure 3-D).

Metabolic profiling by Karasinska et al. [15] subtypes classified most high MG-GEM tumors as glycolytic (75% vs. 25%, p=0.004), and low MG-GEM tumors as cholesterogenic (78% vs. 22%, p=0.004) (Supplementary Figure 6-A). MG-GEM expression correlated with glycolytic programs in the high group (R^2=0.51, p<0.0001) and with cholesterogenic programs in the low group (R^2=0.53, p<0.0001) (Supplementary Figure 6-B). Consistently, “REACTOME_GLYCOLYSIS” was enriched in high MG-GEM tumors (Supplementary Figure 6-C).

Gene Ontology (GO) analysis showed that high MG-GEM tumors were enriched for hypoxia, adhesion, collagen, and apoptosis, while low MG-GEM tumors were enriched for ion channel and lipid antigen pathways (Supplementary Figure 5-B). Reactome confirmed these differences: high MG-GEM tumors were enriched in glucose metabolism, hypoxia, mitochondrial activity, and ROS pathways, whereas low MG-GEM tumors were enriched in lipid homeostasis and channel pathways (Supplementary Figure 7-A), confirming the molecular differences among groups.

Immune-related differences were also analyzed. High MG-GEM tumors showed enrichment in innate immune pathways, while low MG-GEM tumors were enriched in adaptive immune signatures (Supplementary Figure 7-B). MCP-counter deconvolution further revealed reduced immune infiltration in high MG-GEM tumors. Consistently, Myeloid dendritic cells, B cells, T cells (including CD8+), were significantly more abundant in the low MG-GEM group (Supplementary Figure 5-A).

MG-GEM predicts survival in patients who have received adjuvant GEM

To evaluate the prognostic value of MG-GEM, we tested it across independent cohorts. In the remaining Puleo samples (n=211), originally excluded for being classified as intermediate MG stress, Kaplan–Meier analysis confirmed shorter OS (p<0.001) and DFS (p<0.001) in high MG-GEM tumors (n=106) (Figure 4-A) (Figure 4-C). Validation in the Erasme University Hospital cohort (n=42) composed of a series of GEM AT patients with resected PDAC, showed consistent results, with reduced OS (p=0.0025) and DFS (p=0.05) in high MG-GEM patients (Figure 4-D) (Figure 4-F).

We next examined the predictive relevance of MG-GEM for GEM response.. Low MG-GEM tumors expressed higher levels of concentrative nucleoside transporter 1 (CNT1) (p=0.0099), a key transporter for GEM uptake, whereas high MG-GEM tumors showed upregulation of ribonucleotide reductase subunit M1 (RRM1) (p=0.047) and cytidine deaminase (CDA) (p=4.9e-05), both associated with GEM resistance (Supplementary Figure 8).

In the Puleo validation cohort, MG-GEM stratified survival only among patients receiving adjuvant GEM (n = 113; Figure 5A), with low MG-GEM patients showing

longer overall survival (OS; p = 0.009; Figure 5B) and disease-free survival (DFS; p = 0.012; Figure 5C). No significant differences were observed among patients who did not receive adjuvant therapy (n = 71).Finally, in the PRODIGE-24/TGCC PA6 cohort (183 FOLFIRINOX, 167 GEM) (Figure 5-D), MG-GEM predicted outcome only in GEM-treated patients (OS p=0.0017; DFS p=0.0038) , but not in those treated with FOLFIRINOX (OS p=0.66; DFS p=0.82) (Figure 5-E) (Figure 5-F). These results indicate that MG-GEM is a prognostic biomarker that specifically predicts survival benefit following adjuvant GEM treatment.

Independent and Complementary Predictive Value of MG-GEM and GemPred Signatures

We next compared MG-GEM with the GemPred transcriptomic signature (Figure 6-A) [38]. The two signatures were independent (OR=0.698, p=0.314) (Figure 6-B), allowing their combination. Within GEM AT samples, MG-GEM stratified OS in both GemPred+ (p=0.029) (Figure 6-C) and GemPred− (p=0.041) groups (Figure 6-D). However, no significant DFS differences were observed (Supplementary Figure 10), indicating that the combined prognostic value of MG-GEM and GemPred is stronger for OS than DFS.

Discussion

Several transcriptomic signatures have been developed to predict GEM response in PDAC, including GemPred, GemCore, and PancreasView, each providing complementary insights into tumor biology and chemosensitivity. In this study, we identified differentially expressed genes between PDAC tumors with high and low MG scores that most strongly impacted prognosis, leading to the characterization of a novel 16-gene MG stress signature termed MG-GEM. Given the established connection between MG stress and GEM resistance, we investigated whether MG-GEM signature, could serve as a predictive tool for survival benefit to adjuvant GEM therapy and as a prognostic marker in patients with resected PDAC. Stratification into high and low MG-GEM groups identified patient populations with markedly different prognoses, which was validated across two internal and one external independent cohort, the PRODIGE24/CCTG PA6 trial cohort.

Our MG-GEM signature adds a metabolically informed perspective, capturing glycolytic stress and KRAS-driven rewiring that is not fully represented in these existing tools, and represents, to the best of our knowledge, the shortest transcriptomic signature identified to date for predicting GEM response.. Comparison with the GemPred signature [11], demonstrated that MG-GEM provides independent predictive information. Moreover, combining MG-GEM with GemPred improved the identification

of GEM-sensitive and GEM-resistant patients, supporting a complementary role for metabolically informed and clinically derived predictors in personalized treatment strategies. Future clinical translation will require standardized RNA-based assays, consideration of tissue quality and availability, and rapid turnaround to enable timely stratification of patients for adjuvant or neoadjuvant therapies.

Transcriptomic comparison with established PDAC classifications [21], [36], [40], [41], [42], [43], [44], revealed that high MG-GEM tumors were significantly enriched in “Basal-like” and active stroma subtypes, both linked to poor prognosis. In contrast, low MG-GEM tumors were enriched in the “Classical” and immune subtypes, which are associated with better outcomes. Interestingly, a subset of “Pure Classical” tumors fell into the high MG-GEM group, identifying patients who may experience poorer outcomes despite their classical subtype classification. These findings highlight the potential of MG-GEM to refine patient stratification and inform treatment decisions beyond existing classical/basal-like molecular subtyping.

Beyond transcriptomic signatures, additional biomarkers have been explored to guide adjuvant therapy in PDAC. For example, immunohistochemical (IHC) assessment of GATA6, validated in large ESPAC cohorts, has been shown to identify classical versus basal-like subtypes with prognostic and predictive implications [39]. The MG-GEM signature integrates transcriptomic data to reflect functional tumor biology, offering a potential complement to existing IHC and molecular markers.

In PDAC, KRAS mutations contribute to immune evasion and metabolic dysregulation, both of which are key drivers of therapeutic resistance and disease progression [45], [46]. Molecular profiling of MG-GEM-defined subgroups revealed associations with KRAS and SMAD4 mutations. Specifically, KRAS G12D variant was predominantly enriched in the high MG-GEM group, while KRAS G12V was more frequently observed in the low group. Interestingly, a large-scale study investigating the molecular features and clinical outcomes of PDAC patients stratified by KRAS variant revealed that KRAS G12D tumors exhibited a distinct molecular profile, including elevated expression of genes involved in glucose and glutamine metabolism [47].
This suggests that the MG-GEM signature may capture functional heterogeneity downstream of specific KRAS variants. In addition, SMAD4 alterations, which have been implicated in promoting glycolysis [48], and conferring drug resistance through enhanced lipid accumulation [49] were also significantly enriched in the high MG-GEM group. Taken together, these mutational patterns not only provide mechanistic support for the link between MG metabolism and therapeutic outcome but also reinforce the biological plausibility of the MG-GEM signature as a clinically relevant marker of GEM resistance.

Clinically, MG-GEM stratification revealed significant differences in OS and DFS among GEM-treated patients, but not among those treated with FFX, suggesting its specificity for predicting survival benefit from GEM.

Our demonstration of a connection between acquired GEM resistance and MG stress highlighted the involvement of the heat shock response in preclinical PDAC models [20]. Nonetheless, we do not exclude the contribution of additional MG stress-driven mechanisms underlying GEM resistance in glycolytic pancreatic cancer cells. Here, we observed that tumors classified as high MG-GEM exhibited elevated expression of GEM metabolic blockers [20], including CDA and RRM1, compared to low MG-GEM tumors. This expression pattern is consistent with a chemoresistant phenotype and suggests a molecular mechanism linked to the interference with GEM efficacy, warranting further investigation.

This study has limitations, including its reliance on retrospective cohorts and incomplete clinical annotation in some datasets. Future research integrating MG-GEM with comprehensive molecular profiling or predictive biomarkers, as well as evaluating its performance in the neoadjuvant setting, will be important to establish its clinical utility.

In conclusion, we present a novel, biologically grounded MG-GEM signature that captures glycolytic stress–related mechanisms of chemoresistance in PDAC and is associated with prognostic molecular, immune and metabolic features. Validated across independent cohorts, this 16-gene model classifies tumors according to survival benefit from GEM therapy and provides information beyond established molecular subtypes. Importantly, this novel signature warrants further investigation as a potential clinical tool for guiding chemotherapy decisions and opens avenues for exploring KRAS-targeted combination strategies in pancreatic cancer.

Tables

Table 1

Variable	N	Overall N = 85[1]	MG-GEM High Risk N = 43[1]	MG-GEM Low Risk N = 42[1]	p-value[2]
Age	85				0.23
<= 65 years		35 (41%)	15 (35%)	20 (48%)	
>65 years		50 (59%)	28 (65%)	22 (52%)	
Sex	85				0.72
Female		34 (40%)	18 (42%)	16 (38%)	
Male		51 (60%)	25 (58%)	26 (62%)	
Histological grade	83				0.091
G1		27 (33%)	9 (21%)	18 (44%)	
G2		42 (51%)	25 (60%)	17 (41%)	
G3		14 (17%)	8 (19%)	6 (15%)	
Resection Margin	85				0.28
R0		64 (75%)	30 (70%)	34 (81%)	
R1		19 (22%)	11 (26%)	8 (19%)	
R3		2 (2.4%)	2 (4.7%)	0 (0%)	
TNM.Stage	85				0.64
IA		3 (3.5%)	1 (2.3%)	2 (4.8%)	
IB		4 (4.7%)	1 (2.3%)	3 (7.1%)	
IIA		15 (18%)	7 (16%)	8 (19%)	
IIB		63 (74%)	34 (79%)	29 (69%)	
Lymphatic invasion	82	50 (61%)	23 (56%)	27 (66%)	0.37
Perineural invasion	82	62 (76%)	35 (85%)	27 (66%)	**0.040**
Vascular invasion	82	42 (51%)	20 (49%)	22 (54%)	0.66
Adjuvant Treatment	85				0.43
5FU		2 (2.4%)	1 (2.3%)	1 (2.4%)	
Gemcitabine		54 (64%)	24 (56%)	30 (71%)	

Variable	N	Overall N = 85[1]	MG-GEM High Risk N = 43[1]	MG-GEM Low Risk N = 42[1]	p-value[2]
Other		2 (2.4%)	1 (2.3%)	1 (2.4%)	
Untreated		27 (32%)	17 (40%)	10 (24%)	

[1]Median (IQR) or Frequency (%)

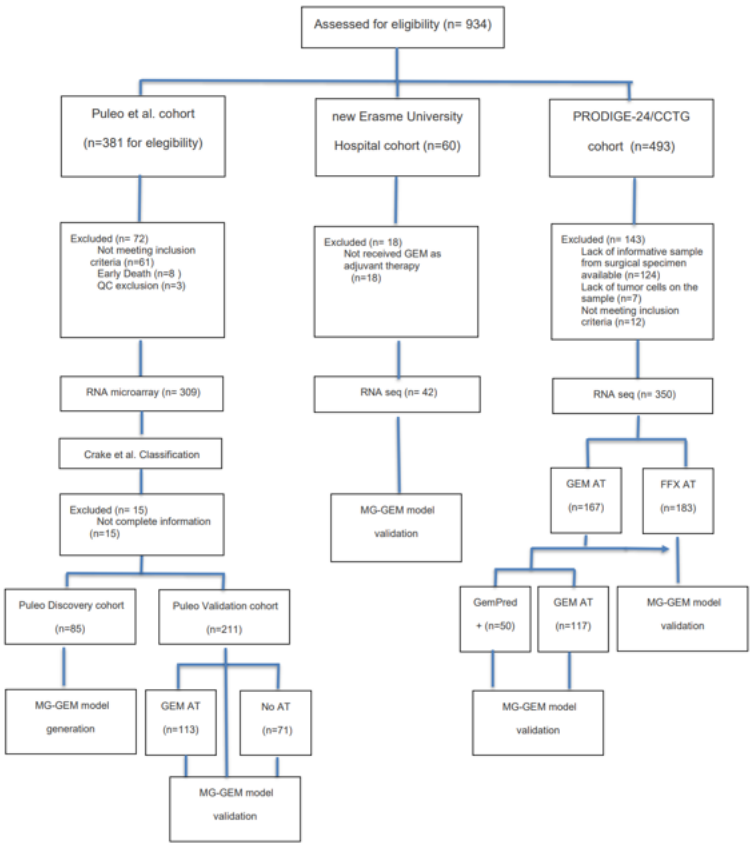
[2]Pearson's Chi-squared test; Fisher's exact test

Table1. Clinical Characteristics of the high and low MG-GEM patient groups.

Figures

Figure 1

A)



B)

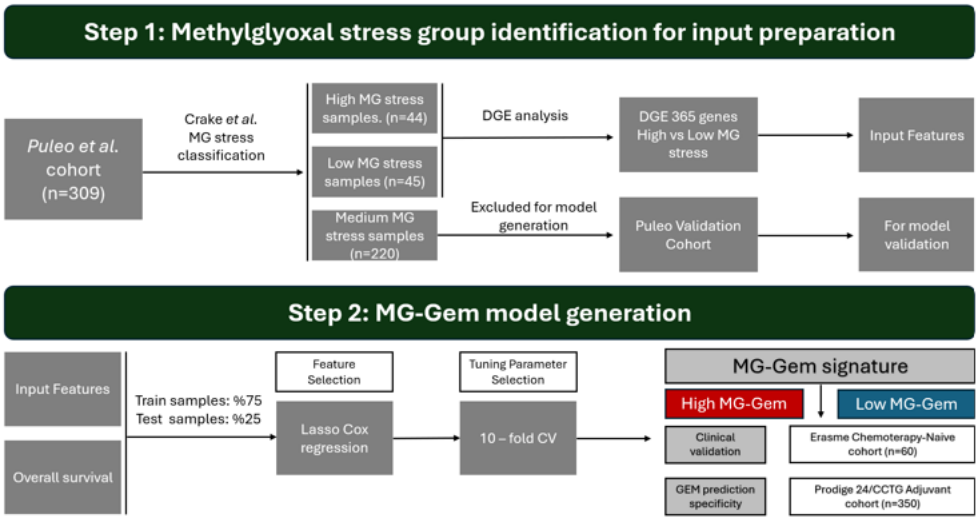


Figure 1. Workflow of the study.

(A) CONSORT diagram of the procedure followed in the work.
(B) Overview of the study workflow.

Abbreviations: MG, Methylglyoxal; DGE, Differential Gene Expression; CV, Cross-Validation; GEM, Gemcitabine.

Figure 2

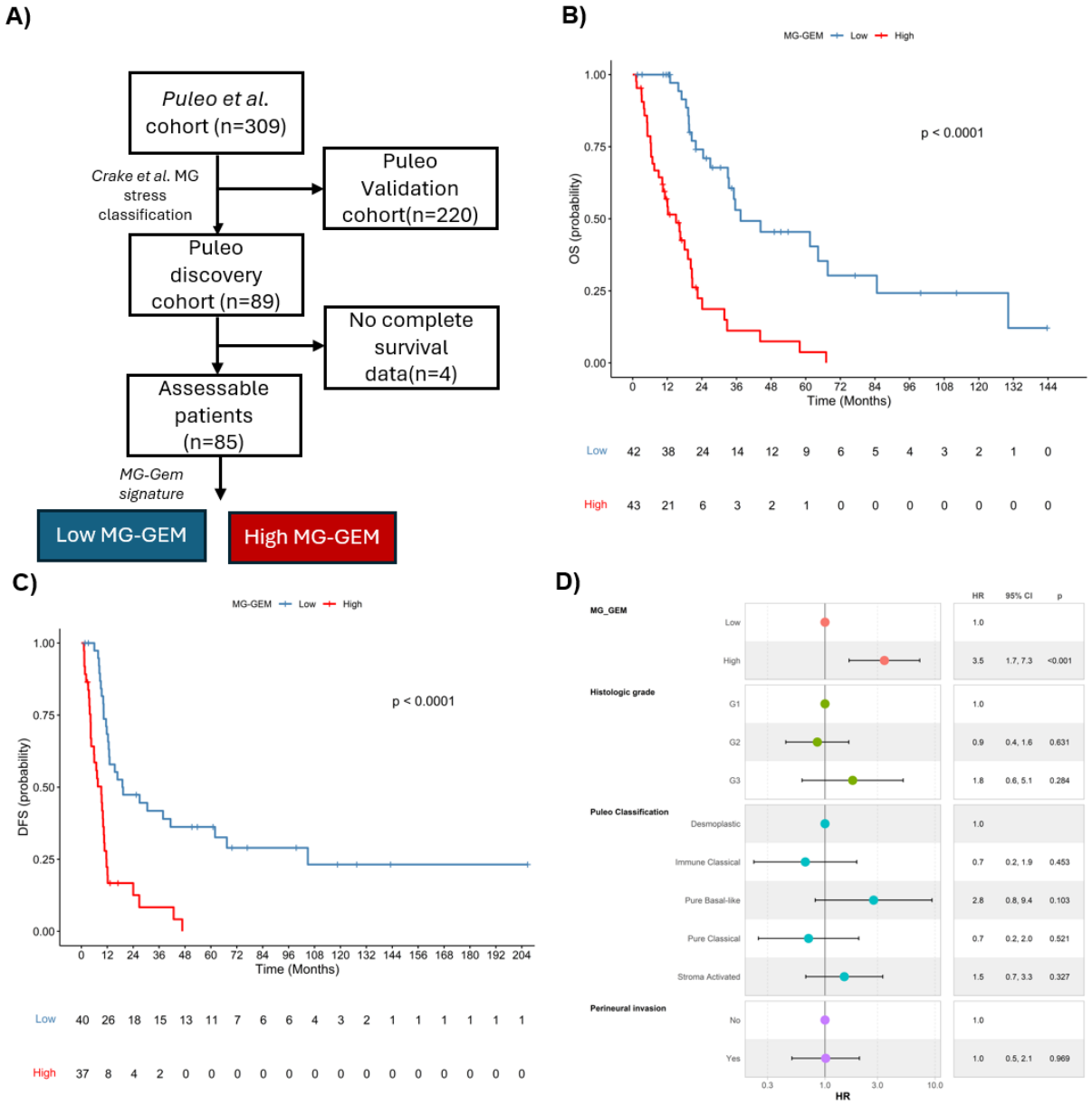

A)
Puleo et al. cohort (n=309)
Crake et al. MG stress classification
Puleo Validation cohort(n=220)
Puleo discovery cohort (n=89)
No complete survival data(n=4)
Assessable patients (n=85)
MG-Gem signature
Low MG-GEM
High MG-GEM
B)
MG-GEM Low High
p < 0.0001
OS (probability)
Time (Months)
Low 42 38 24 14 12 9 6 5 4 3 2 1 0
High 43 21 6 3 2 1 0 0 0 0 0 0 0
C)
MG-GEM Low High
p < 0.0001
DFS (probability)
Time (Months)
Low 40 26 18 15 13 11 7 6 6 4 3 2 1 1 1 1 1
High 37 8 4 2 0 0 0 0 0 0 0 0 0 0 0 0 0
D)
HR 95% CI p
MG_GEM
Low 1.0
High 3.5 1.7, 7.3 <0.001
Histologic grade
G1 1.0
G2 0.9 0.4, 1.6 0.631
G3 1.8 0.6, 5.1 0.284
Puleo Classification
Desmoplastic 1.0
Immune Classical 0.7 0.2, 1.9 0.453
Pure Basal-like 2.8 0.8, 9.4 0.103
Pure Classical 0.7 0.2, 2.0 0.521
Stroma Activated 1.5 0.7, 3.3 0.327
Perineural invasion
No 1.0
Yes 1.0 0.5, 2.1 0.969
HR

Figure 2. Survival analysis of MG-GEM stratification in the Puleo discovery cohort (n = 89)
(A) Flowchart of the discovery cohort.
(B, C) Kaplan-Meier analysis of overall survival (OS) and disease-free survival (DFS).
(D) Forest plot showing multivariate Cox proportional hazards regression including MG-GEM groups.

Abbreviations: MG, Methylglyoxal; GEM, Gemcitabine; OS, Overall Survival; DFS, Disease-Free Survival; HR, Hazard Ratio; CI, Confidence Interval.

Figure 3

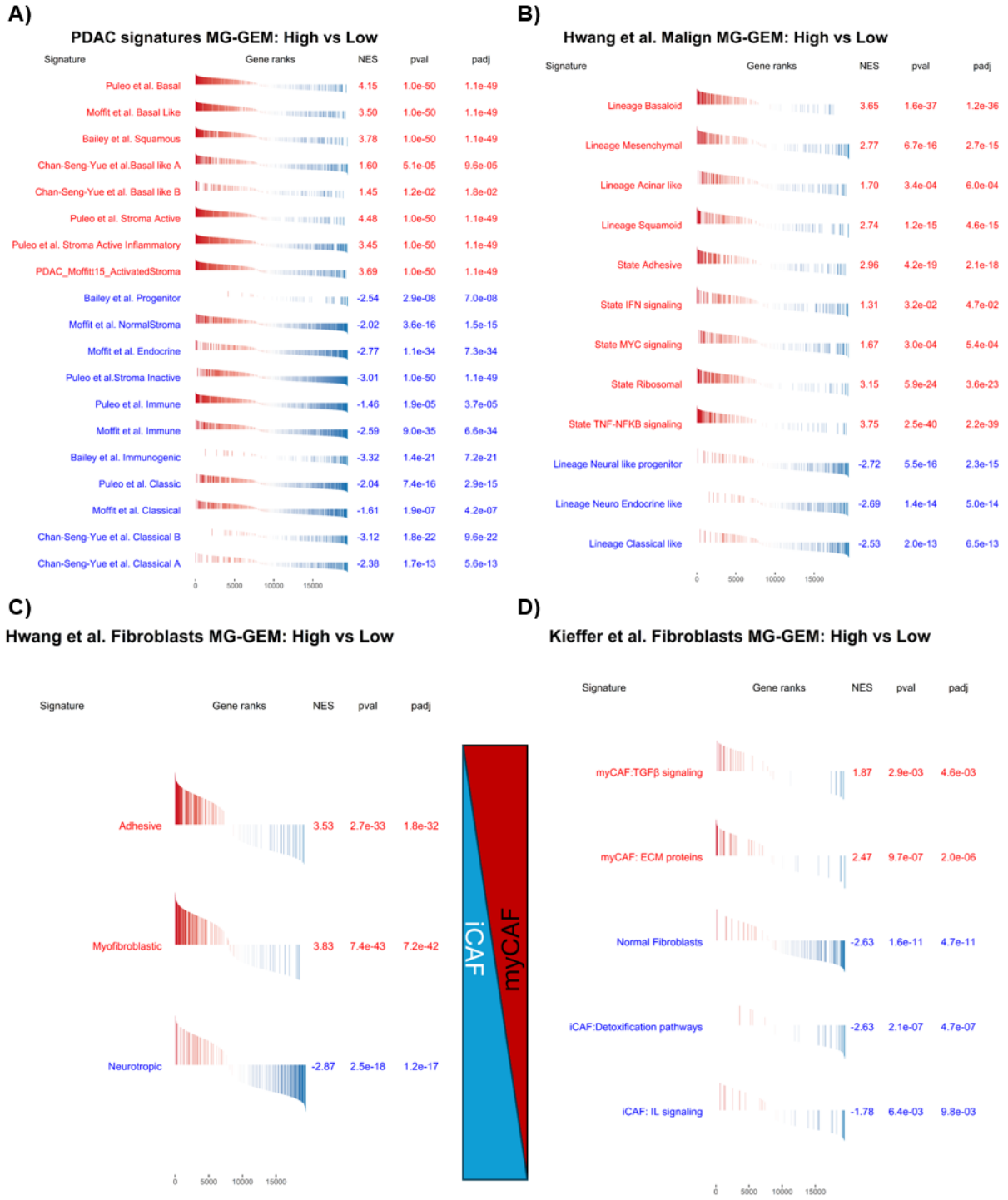

A)
PDAC signatures MG-GEM: High vs Low
Signature
Gene ranks
NES
pval
padj
Puleo et al. Basal 4.15 1.0e-50 1.1e-49
Moffit et al. Basal Like 3.50 1.0e-50 1.1e-49
Bailey et al. Squamous 3.78 1.0e-50 1.1e-49
Chan-Seng-Yue et al.Basal like A 1.60 5.1e-05 9.6e-05
Chan-Seng-Yue et al. Basal like B 1.45 1.2e-02 1.8e-02
Puleo et al. Stroma Active 4.48 1.0e-50 1.1e-49
Puleo et al. Stroma Active Inflammatory 3.45 1.0e-50 1.1e-49
PDAC_Moffitt15_ActivatedStroma 3.69 1.0e-50 1.1e-49
Bailey et al. Progenitor -2.54 2.9e-08 7.0e-08
Moffit et al. NormalStroma -2.02 3.6e-16 1.5e-15
Moffit et al. Endocrine -2.77 1.1e-34 7.3e-34
Puleo et al.Stroma Inactive -3.01 1.0e-50 1.1e-49
Puleo et al. Immune -1.46 1.9e-05 3.7e-05
Moffit et al. Immune -2.59 9.0e-35 6.6e-34
Bailey et al. Immunogenic -3.32 1.4e-21 7.2e-21
Puleo et al. Classic -2.04 7.4e-16 2.9e-15
Moffit et al. Classical -1.61 1.9e-07 4.2e-07
Chan-Seng-Yue et al. Classical B -3.12 1.8e-22 9.6e-22
Chan-Seng-Yue et al. Classical A -2.38 1.7e-13 5.6e-13
0 5000 10000 15000
B)
Hwang et al. Malign MG-GEM: High vs Low
Signature
Gene ranks
NES
pval
padj
Lineage Basaloid 3.65 1.6e-37 1.2e-36
Lineage Mesenchymal 2.77 6.7e-16 2.7e-15
Lineage Acinar like 1.70 3.4e-04 6.0e-04
Lineage Squamoid 2.74 1.2e-15 4.6e-15
State Adhesive 2.96 4.2e-19 2.1e-18
State IFN signaling 1.31 3.2e-02 4.7e-02
State MYC signaling 1.67 3.0e-04 5.4e-04
State Ribosomal 3.15 5.9e-24 3.6e-23
State TNF-NFKB signaling 3.75 2.5e-40 2.2e-39
Lineage Neural like progenitor -2.72 5.5e-16 2.3e-15
Lineage Neuro Endocrine like -2.69 1.4e-14 5.0e-14
Lineage Classical like -2.53 2.0e-13 6.5e-13
0 5000 10000 15000
C)
Hwang et al. Fibroblasts MG-GEM: High vs Low
Signature
Gene ranks
NES
pval
padj
Adhesive 3.53 2.7e-33 1.8e-32
Myofibroblastic 3.83 7.4e-43 7.2e-42
Neurotropic -2.87 2.5e-18 1.2e-17
0 5000 10000 15000
iCAF
myCAF
D)
Kieffer et al. Fibroblasts MG-GEM: High vs Low
Signature
Gene ranks
NES
pval
padj
myCAF:TGFβ signaling 1.87 2.9e-03 4.6e-03
myCAF: ECM proteins 2.47 9.7e-07 2.0e-06
Normal Fibroblasts -2.63 1.6e-11 4.7e-11
iCAF:Detoxification pathways -2.63 2.1e-07 4.7e-07
iCAF: IL signaling -1.78 6.4e-03 9.8e-03
0 5000 10000 15000

Figure 3. Association of high and low MG-GEM samples with major PDAC molecular subtypes from bulk and single-cell transcriptomic analyses

(A) Normalized enrichment scores (NES) showing that the high MG-GEM group is significantly enriched for “Basal-like” and “Stroma Active” signatures from bulk molecular classifications.
(B) NES indicating that the high MG-GEM group is significantly enriched in “Basaloid,” “Mesenchymal,” and “Squamoid” single-cell–derived signatures from Hwang et al.
(C, D) NES from GSEA of CAF signatures (Hwang et al. and Kieffer et al.) showing enrichment of myCAFs in high MG-GEM samples and enrichment of iCAFs in low MG-GEM samples.

Figure 4

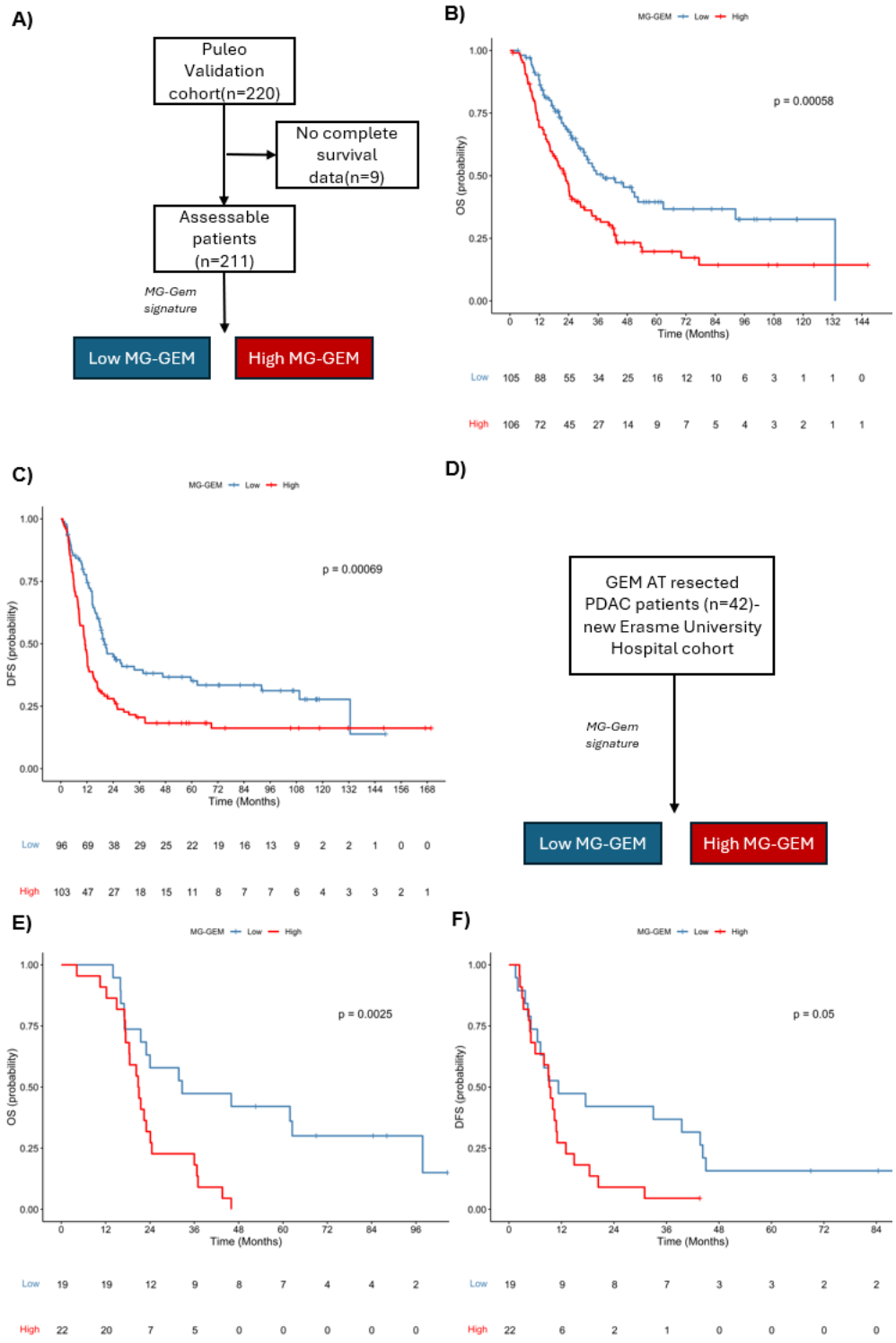

A)
Puleo Validation cohort(n=220)
No complete survival data(n=9)
Assessable patients (n=211)
MG-Gem signature
Low MG-GEM
High MG-GEM
B)
MG-GEM Low High
p = 0.00058
OS (probability)
Time (Months)
Low 105 88 55 34 25 16 12 10 6 3 1 1 0
High 106 72 45 27 14 9 7 5 4 3 2 1 1
C)
MG-GEM Low High
p = 0.00069
DFS (probability)
Time (Months)
Low 96 69 38 29 25 22 19 16 13 9 2 2 1 0 0
High 103 47 27 18 15 11 8 7 7 6 4 3 3 2 1
D)
GEM AT resected PDAC patients (n=42)- new Erasme University Hospital cohort
MG-Gem signature
Low MG-GEM
High MG-GEM
E)
MG-GEM Low High
p = 0.0025
OS (probability)
Time (Months)
Low 19 19 12 9 8 7 4 4 2
High 22 20 7 5 0 0 0 0 0
F)
MG-GEM Low High
p = 0.05
DFS (probability)
Time (Months)
Low 19 9 8 7 3 3 2 2
High 22 6 2 1 0 0 0 0

Figure 4. Survival analyses validating the prognostic value of MG-GEM stratification

(A) Flowchart of the Puleo validation cohort (n = 220).
(B, C) Kaplan-Meier analysis of the Puleo validation cohort showing significantly worse overall survival (OS).
(D) Flowchart of GEM-adjuvant resected PDAC patients from the Erasme University Hospital cohort (n = 42).
(E, F) Kaplan-Meier analysis of the chemotherapy-naive Erasme University Hospital cohort showing significantly worse OS and DFS in the high MG-GEM group.

Abbreviations: MG, Methylglyoxal; OS, Overall Survival; DFS, Disease-Free Survival.

Figure 5

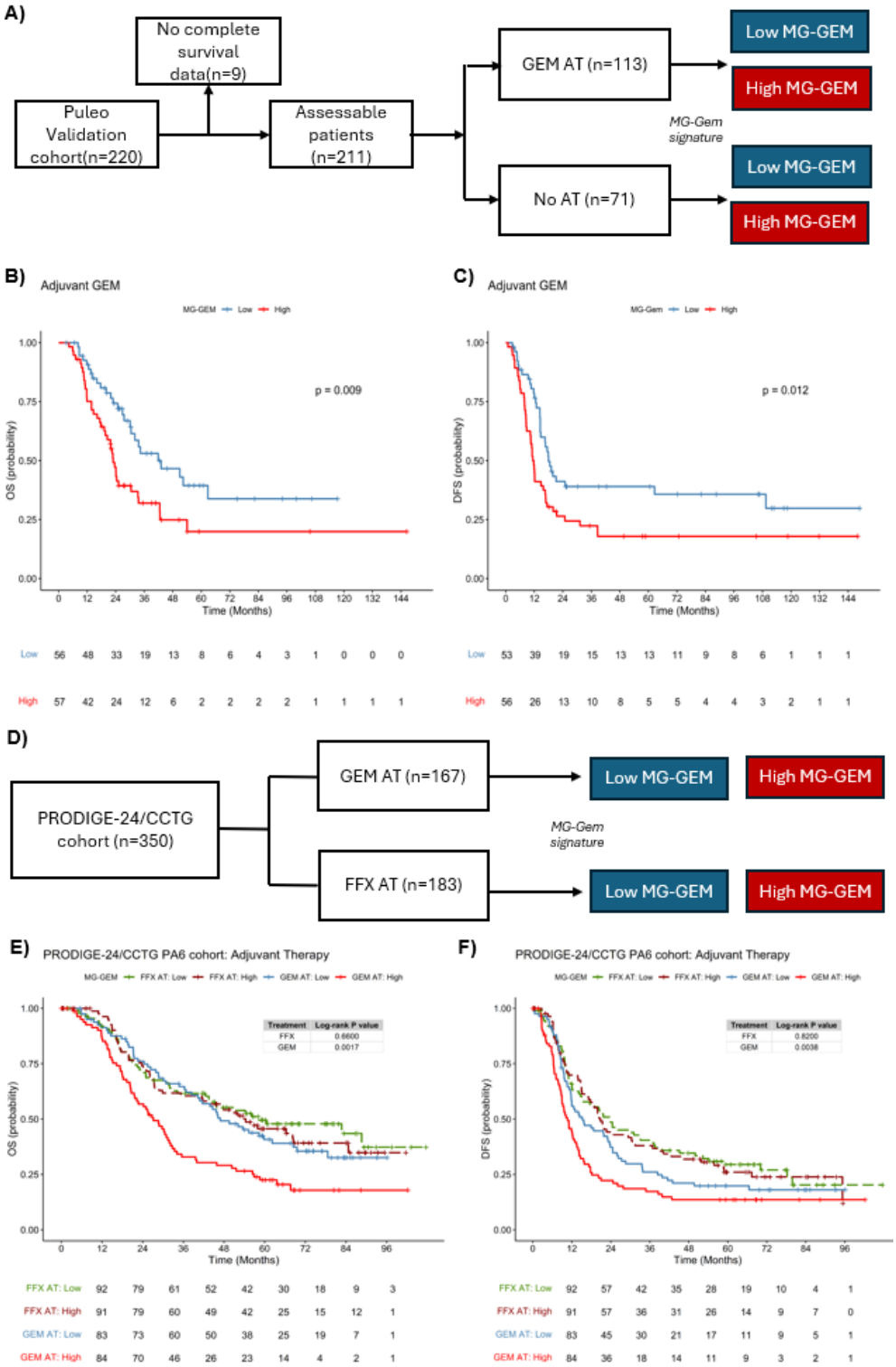

A)
No complete survival data(n=9)
Puleo Validation cohort(n=220)
Assessable patients (n=211)
GEM AT (n=113)
No AT (n=71)
MG-Gem signature
Low MG-GEM
High MG-GEM
Low MG-GEM
High MG-GEM
B) Adjuvant GEM
MG-GEM Low High
p = 0.009
OS (probability)
Time (Months)
Low 56 48 33 19 13 8 6 4 3 1 0 0 0
High 57 42 24 12 6 2 2 2 2 1 1 1 1
C) Adjuvant GEM
MG-Gem Low High
p = 0.012
DFS (probability)
Time (Months)
Low 53 39 19 15 13 13 11 9 8 6 1 1 1
High 56 26 13 10 8 5 5 4 4 3 2 1 1
D)
PRODIGE-24/CCTG cohort (n=350)
GEM AT (n=167)
FFX AT (n=183)
MG-Gem signature
Low MG-GEM
High MG-GEM
Low MG-GEM
High MG-GEM
E) PRODIGE-24/CCTG PA6 cohort: Adjuvant Therapy
MG-GEM FFX AT: Low FFX AT: High GEM AT: Low GEM AT: High
Treatment Log-rank P value
FFX 0.6600
GEM 0.0017
OS (probability)
Time (Months)
FFX AT: Low 92 79 61 52 42 30 18 9 3
FFX AT: High 91 79 60 49 42 25 15 12 1
GEM AT: Low 83 73 60 50 38 25 19 7 1
GEM AT: High 84 70 46 26 23 14 4 2 1
F) PRODIGE-24/CCTG PA6 cohort: Adjuvant Therapy
MG-GEM FFX AT: Low FFX AT: High GEM AT: Low GEM AT: High
Treatment Log-rank P value
FFX 0.8200
GEM 0.0038
DFS (probability)
Time (Months)
FFX AT: Low 92 57 42 35 28 19 10 4 1
FFX AT: High 91 57 36 31 26 14 9 7 0
GEM AT: Low 83 45 30 21 17 11 9 5 1
GEM AT: High 84 36 18 14 11 9 3 2 1

Figure 5. **MG-GEM specifically predicts survival benefit from GEM therapy but not to FFX.**

(A) Puleo Validation cohort flowchart stratified by the type of adjuvant treatment.

(B, C) Kaplan-Meier analysis of the adjuvant gemcitabine–treated Puleo validation cohort showing significantly worse overall survival (OS) and disease-free survival (DFS) in the high MG-GEM group.

(D) PRODIGE 24/CCTG PA6 cohort flowchart.

(E,F) Kaplan-Meier analysis of patients from the PRODIGE 24/CCTG PA6 study. Patients stratified as high MG-GEM and receiving adjuvant gemcitabine (red curve) exhibited significantly poorer overall survival (OS) and disease-free survival (DFS) compared to the low MG-GEM group (blue curve). No significant differences were observed among patients who received adjuvant FOLFIRINOX (brown and green curves).

Abbreviations: MG, Methylglyoxal; OS, Overall Survival; DFS, Disease-Free Survival; AT, Adjuvant Therapy; GEM, Gemcitabine; FFX, FOLFIRINOX.

Figure 6

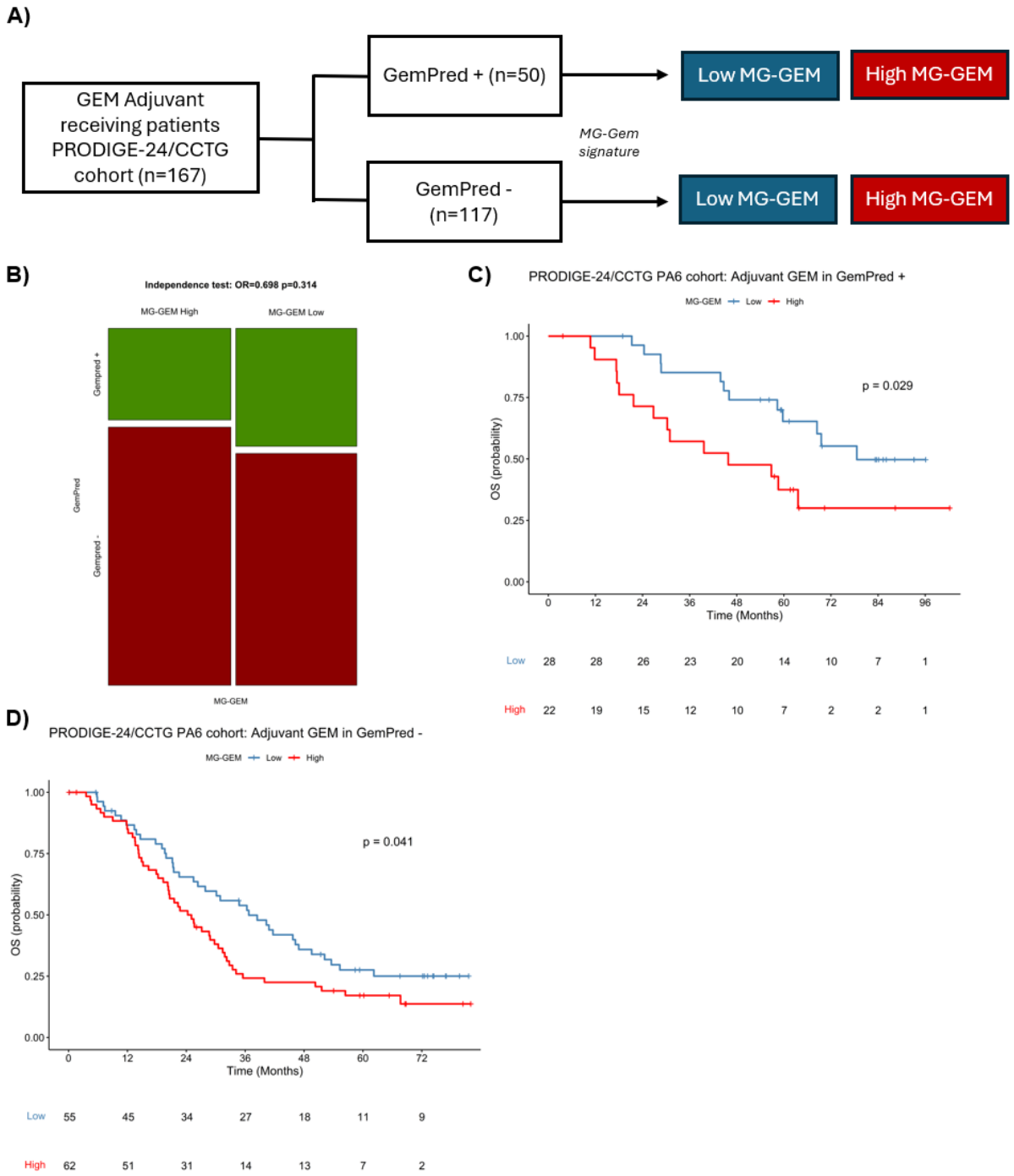

A)
GEM Adjuvant receiving patients PRODIGE-24/CCTG cohort (n=167)
GemPred + (n=50)
GemPred - (n=117)
MG-Gem signature
Low MG-GEM
High MG-GEM
Low MG-GEM
High MG-GEM
B)
Independence test: OR=0.698 p=0.314
MG-GEM High
MG-GEM Low
Gempred +
Gempred -
GemPred
MG-GEM
C)
PRODIGE-24/CCTG PA6 cohort: Adjuvant GEM in GemPred +
MG-GEM Low High
p = 0.029
OS (probability)
Time (Months)
Low 28 28 26 23 20 14 10 7 1
High 22 19 15 12 10 7 2 2 1
D)
PRODIGE-24/CCTG PA6 cohort: Adjuvant GEM in GemPred -
MG-GEM Low High
p = 0.041
OS (probability)
Time (Months)
Low 55 45 34 27 18 11 9
High 62 51 31 14 13 7 2

Figure 6. Independent Validation and Integration of MG-GEM and GemPred Signatures

(A) Individual flowcharts of the GEM-adjuvant PRODIGE-24/CCTG PA6 cohort stratified by GemPred groups.

(B) Fisher's exact test showing independence between MG-GEM and GemPred groups among samples treated with adjuvant gemcitabine in the PRODIGE-24/CCTG PA6 cohort.

(C, D) Kaplan-Meier analysis of the adjuvant gemcitabine–treated PRODIGE-24/CCTG PA6 cohort showing significantly worse overall survival (OS) in the high MG-GEM group within the GemPred+ (C) and GemPred– (D) subgroups.

Abbreviations: OS; Overall Survival; MG: Methylglyoxal; AT: Adjuvant therapy; GEM: Gemcitabine.

Supplementary Tables

Suppl Table 1

Genes	Regression Coefficients
FAM162A	0,0125
SEC61G	0,0122
NT5E	0,0619
TREM1	0,0069
SLC3A1	-0,1342
CALB2	0,1638
TFF1	-0,0203
SLC11A1	0,3106
PDIA2	0,3100
PADI1	0,2065
ADH1C	-0,0441
FXYD3	-0,2561
SCEL	0,2446
ECT2	-0,0466
TIPARP	0,0141
IGHG1	-0,1351

Supplementary table 1. LASSO Cox regression coefficients of the 16 genes used, together with their expression levels, to calculate the final MG-GEM risk score.

Suppl Table 2

			MG-GEM		
Variable	**N**	**Overall** N = 85[1]	**High Risk** N = 43[1]	**Low Risk** N = 42[1]	**p-value**[2]
***KRAS* mutation**	62				**0.043**
G12D		29 (47%)	19 (59%)	10 (33%)	
G12V		23 (37%)	7 (22%)	16 (53%)	
wt		10 (16%)	6 (19%)	4 (13%)	
***TP53* mutation**	78				0.20
nonsense		47 (60%)	27 (68%)	20 (53%)	
wt		29 (37%)	12 (30%)	17 (45%)	
indel_frameshift		1 (1.3%)	1 (2.5%)	0 (0%)	
indel_inframe		1 (1.3%)	0 (0%)	1 (2.6%)	
***SMAD4* mutation**	78				**0.049**
wt		64 (82%)	29 (73%)	35 (92%)	
indel_frameshift		0 (0%)	0 (0%)	0 (0%)	
nonsense		13 (17%)	10 (25%)	3 (7.9%)	
indel_inframe		1 (1.3%)	1 (2.5%)	0 (0%)	
***CDKN2A* mutation**	78				0.16
wt		63 (81%)	29 (73%)	34 (89%)	
indel_frameshift		6 (7.7%)	5 (13%)	1 (2.6%)	
nonsense		8 (10%)	5 (13%)	3 (7.9%)	
indel_inframe		1 (1.3%)	1 (2.5%)	0 (0%)	
Puleo et al.	77				**<0.001**
Desmoplastic		16 (21%)	10 (25%)	6 (16%)	
Immune Classical		13 (17%)	1 (2.5%)	12 (32%)	
Pure Basal-like		9 (12%)	8 (20%)	1 (2.7%)	
Pure Classical		19 (25%)	6 (15%)	13 (35%)	
Stroma Activated		20 (26%)	15 (38%)	5 (14%)	
Bailey et al.	85				**<0.001**

Variable	N	Overall N = 85[1]	MG-GEM High Risk N = 43[1]	MG-GEM Low Risk N = 42[1]	p-value[2]
ADEX		20 (24%)	8 (19%)	12 (29%)	
Immunogenic		7 (8.2%)	1 (2.3%)	6 (14%)	
Progenitor		25 (29%)	5 (12%)	20 (48%)	
Squamous		33 (39%)	29 (67%)	4 (9.5%)	
Moffit et al. Tumor	85				0.45
Basal		46 (54%)	25 (58%)	21 (50%)	
Classical		39 (46%)	18 (42%)	21 (50%)	
Moffit et al. Stroma	85				**0.014**
Absent		29 (34%)	13 (30%)	16 (38%)	
Activated		47 (55%)	29 (67%)	18 (43%)	
Normal		9 (11%)	1 (2.3%)	8 (19%)	
Purist	77				**0.008**
Basal		16 (21%)	13 (33%)	3 (8.1%)	
Classical		61 (79%)	27 (68%)	34 (92%)	

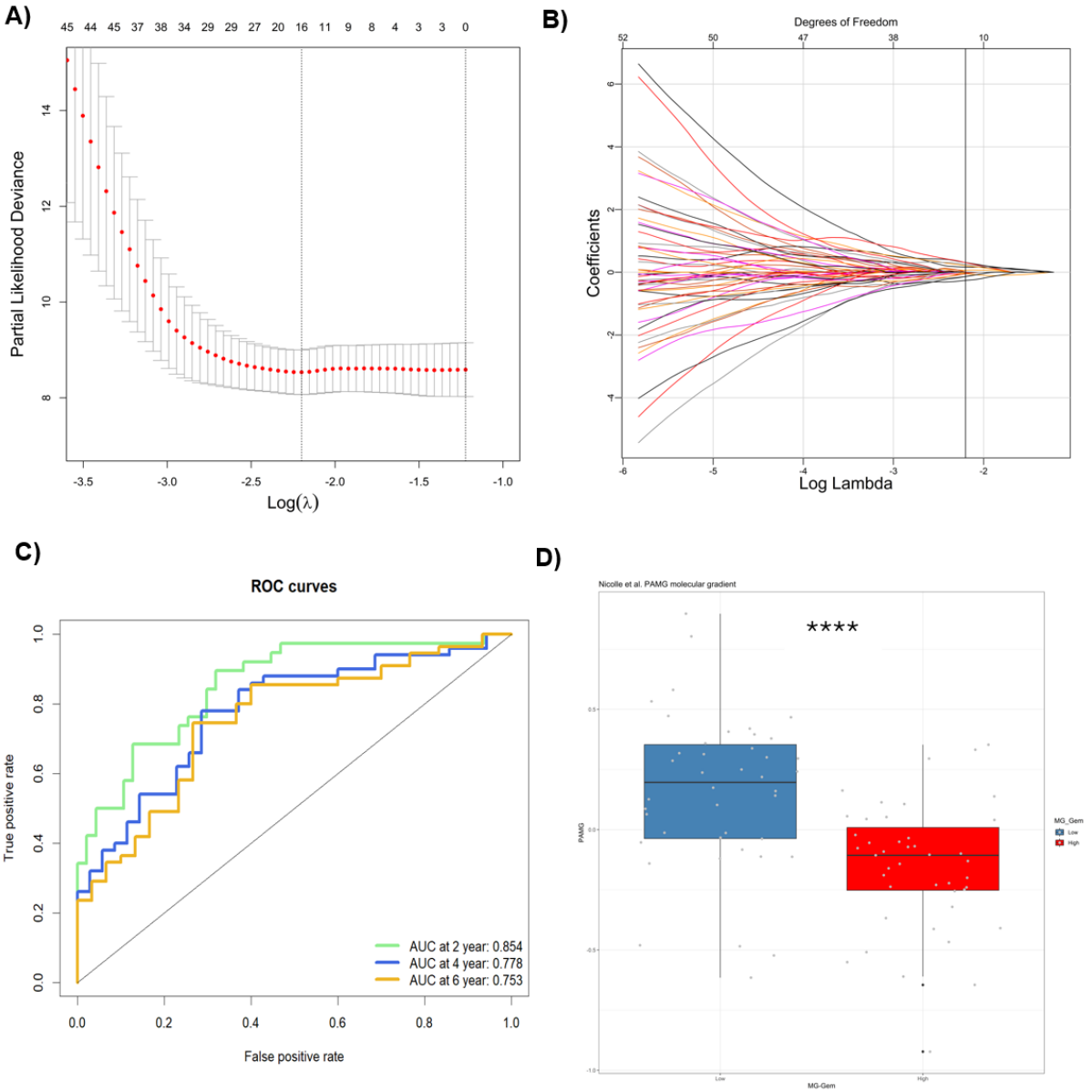
[1]Median (IQR) or Frequency (%)

[2]Fisher's exact test; Pearson's Chi-squared test

Supplementary Table 2. Molecular characteristics of patients stratified by high and low MG-GEM groups.

Supplementary Figures

Supp Figure 1



Supplementary Figure 1. Feature selection and model evaluation using LASSO Cox regression with 10-fold cross-validation

(A) LASSO coefficients from the Cox regression model.

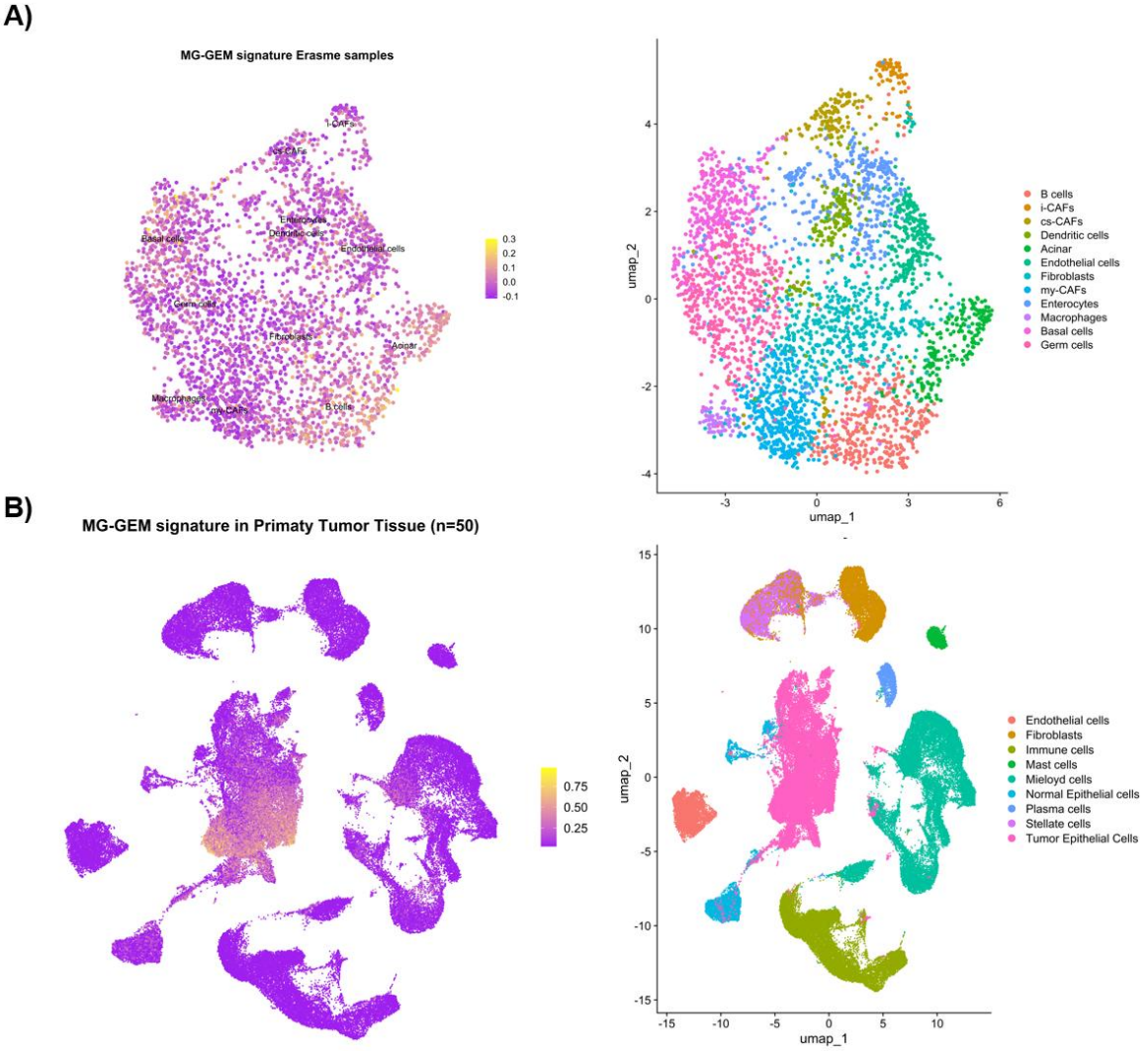
(B) Coefficient profile plot across the log(λ) sequence.

(C) Comparison of ROC curves for the risk score at selected time points.

(D) Continuous PDAC molecular gradient (PAMG) applied in the discovery cohort, **showing significantly higher PAMG scores associated with low MG-GEM samples.**

Abbreviation: AUC, Area Under the Curve.

Supp Figure 2



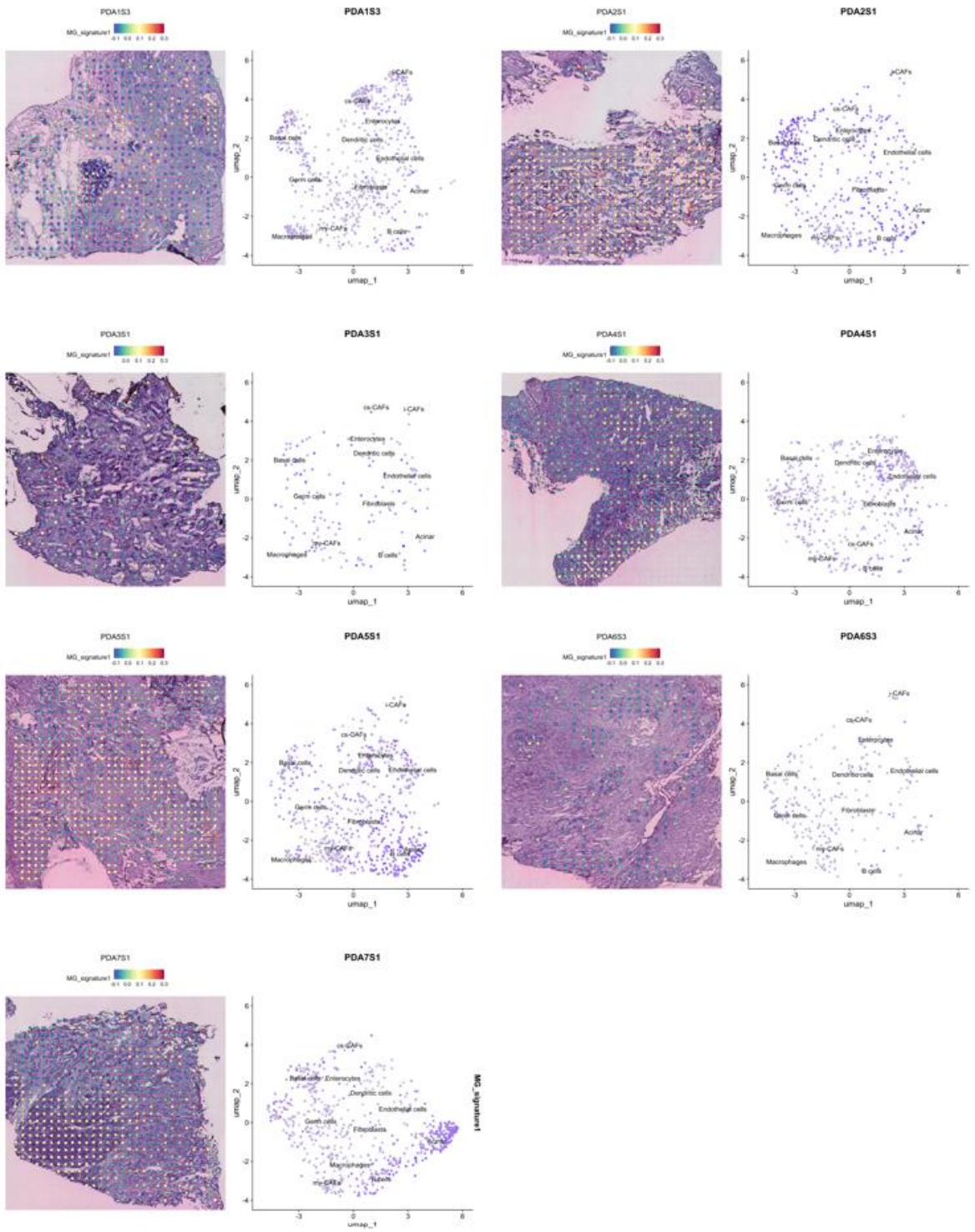
Supplementary Figure 2. Expression of MG-GEM signature genes in basal cells

(A) Feature plots of module scores and cluster annotations showing enrichment of MG-GEM signature genes in acinar and basal cells in Erasme ST samples (n = 7). **(B)** Feature plots of module scores and cluster annotations showing enrichment of signature genes in tumor epithelial cells across 50 primary tumor scRNA-seq datasets.

Abbreviations: MG, Methylglyoxal; ST, Spatial Transcriptomics; scRNA-seq, single-cell RNA sequencing.

Supp Figure 3

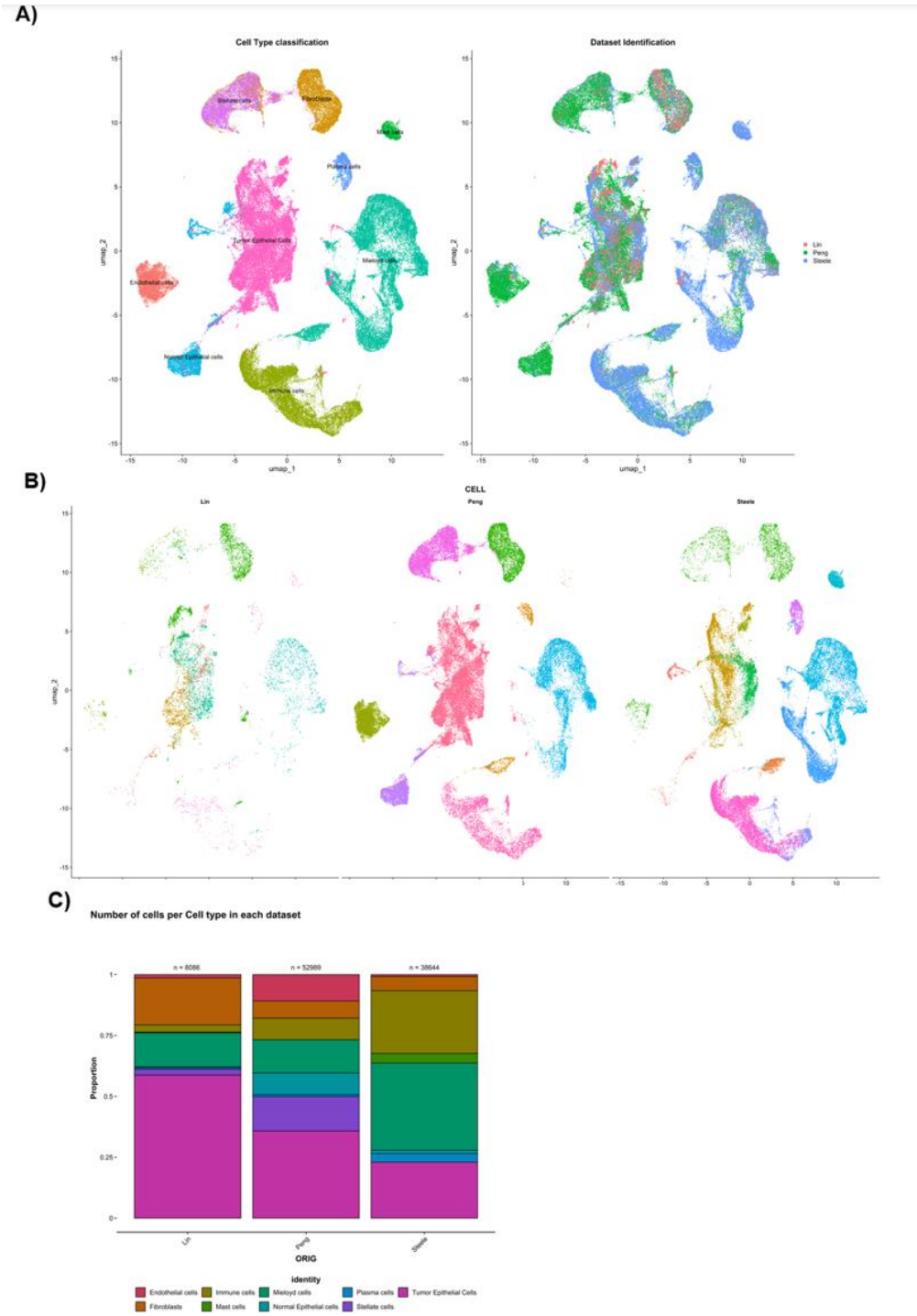
A)



Supplementary Figure 3. Validation of MG-GEM prognostic significance

(A) H&E image and UMAP plots of MG signature gene expression in cohort samples (n = 7).

Supp Figure 4

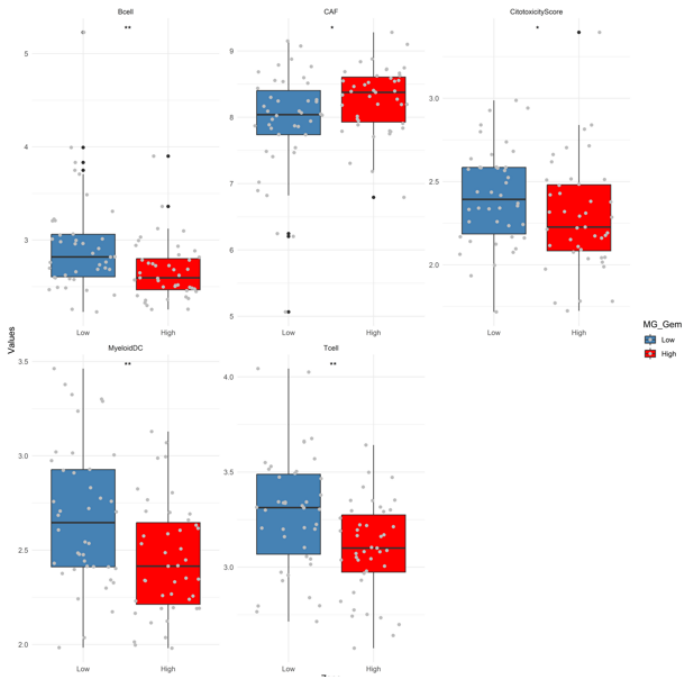


Supplementary Figure 4. Integration of primary tumor scRNA-seq data from three cohorts
(A) UMAP plots displaying cell types and original sample IDs in the integrated dataset.
(B) UMAP plots of each individual dataset showing annotated cell types.
(C) Bar plots depicting the cell composition of each original dataset.

Abbreviation: scRNA-seq, single-cell RNA sequencing.

Supp Figure 5

A)



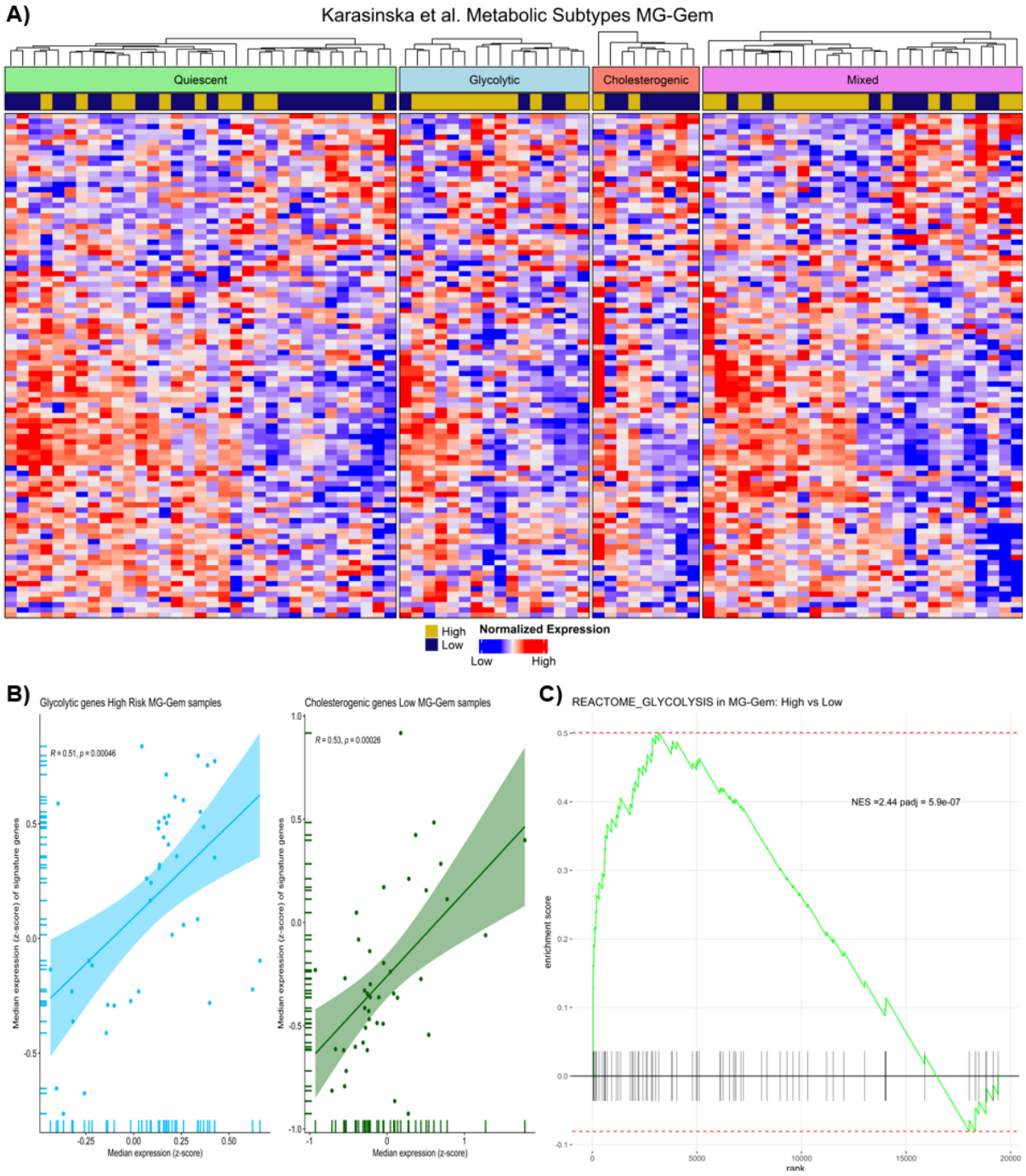
B)

Supplementary Figure 5. Cancer hallmark signature GSEAs

(A) Cell type enrichment analysis using MCPcounter revealed significantly lower scores for cytotoxic cells, T cells, B cells, and myeloid dendritic cells, and higher cancer-associated fibroblast (CAF) scores in high MG-GEM samples.
(B) Normalized enrichment scores (NES) of GO terms related to cancer hallmarks significantly enriched in high MG-GEM versus low MG-GEM samples.

Abbreviations: MG, Methylglyoxal; GSEA, Gene Set Enrichment Analysis; GO, Gene Ontology.

Supp Figure 6



Supplementary Figure 6. Correlation of MG-GEM groups with Karasinska et al. metabolic subtypes

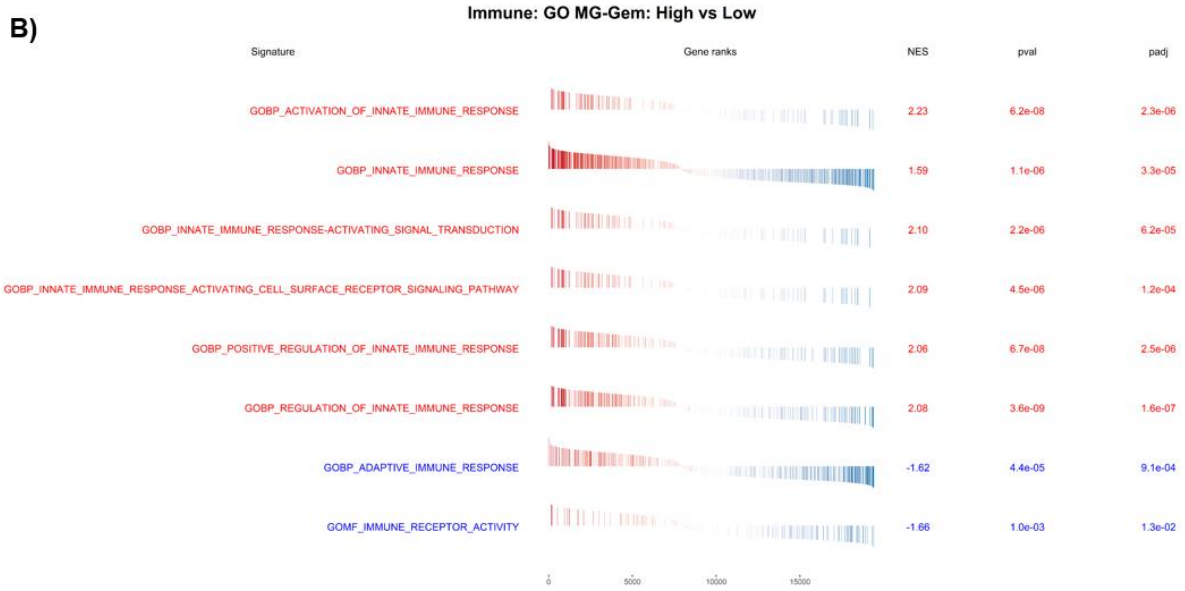
(A) Heatmap showing the characterization of Karasinska et al. metabolic subtypes in the Puleo discovery cohort, with glycolytic samples predominantly in the high MG-GEM group and cholesterogenic samples in the low MG-GEM group.
(B) Positive Pearson correlation between the median expression of glycolytic subtype genes and MG-GEM signature genes in high MG-GEM samples, and a Pearson correlation between cholesterogenic subtype genes and MG-GEM signature genes in low MG-GEM samples.
(C) GSEA results showing positive enrichment of the pathway **REACTOME_GLYCOLYSIS** in high MG-GEM versus low MG-GEM samples.

Abbreviations: MG, Methylglyoxal; GSEA, Gene Set Enrichment Analysis.

Supp Figure 7

A)

B)



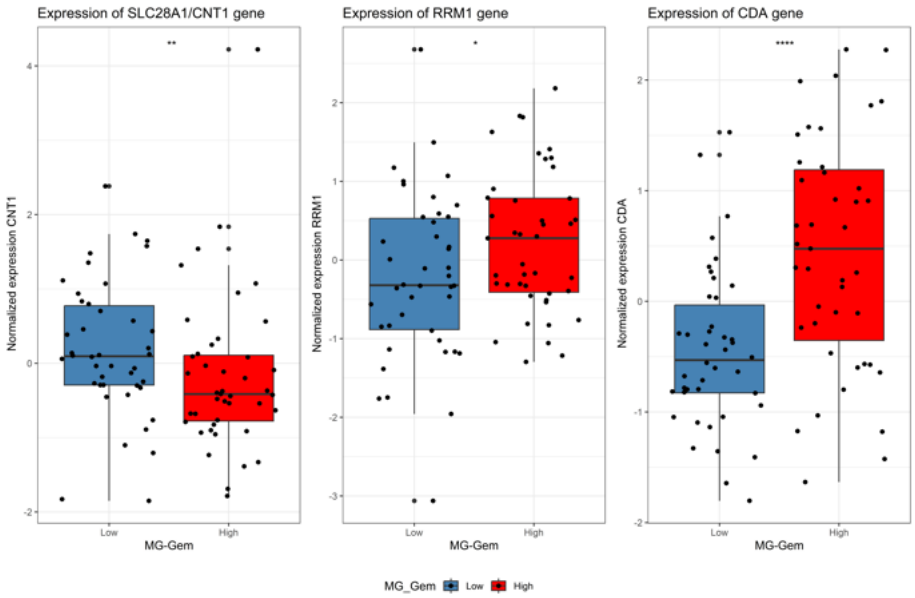
Supplementary Figure 7. Reactome and immune-related GO term enrichment (GSEA)

(A) Normalized enrichment scores (NES) of Reactome pathways significantly enriched in the high MG-GEM versus low MG-GEM groups.

(B) NES of immune-related GO terms showing enrichment of innate immunity pathways in the high MG-GEM group and adaptive immunity pathways in the low MG-GEM group.

Abbreviations: MG, Methylglyoxal; GSEA, Gene Set Enrichment Analysis; GO, Gene Ontology.

Supp Figure 8

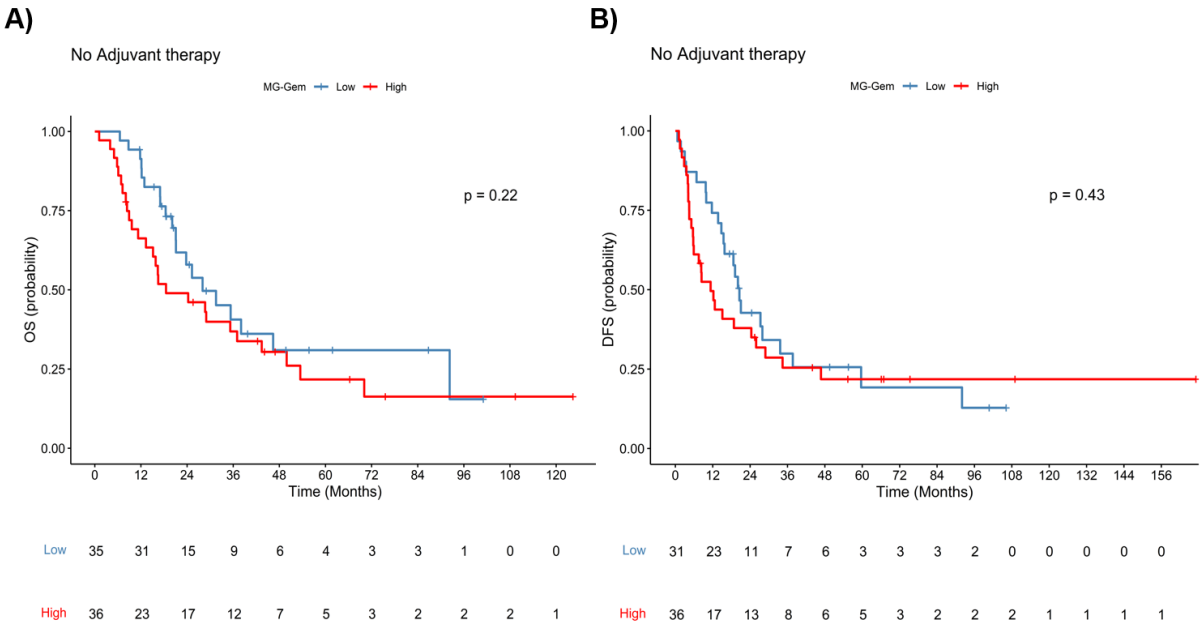


Supplementary Figure 8. Expression of genes involved in gemcitabine (GEM) transport and metabolism in high and low MG-GEM groups

The low MG-GEM group showed higher expression of the GEM transporter CNT1 ($p = 0.0099$), which mediates GEM uptake. In contrast, the high MG-GEM group exhibited elevated expression of two enzymes associated with GEM resistance: RRM1 ($p = 0.047$) and CDA ($p = 4.9 \times 10^{-5}$).

Abbreviations: CNT1, Concentrative Nucleoside Transporter 1; RRM1, Ribonucleotide Reductase Catalytic Subunit M1; CDA, Cytidine Deaminase.

Supp Figure 9

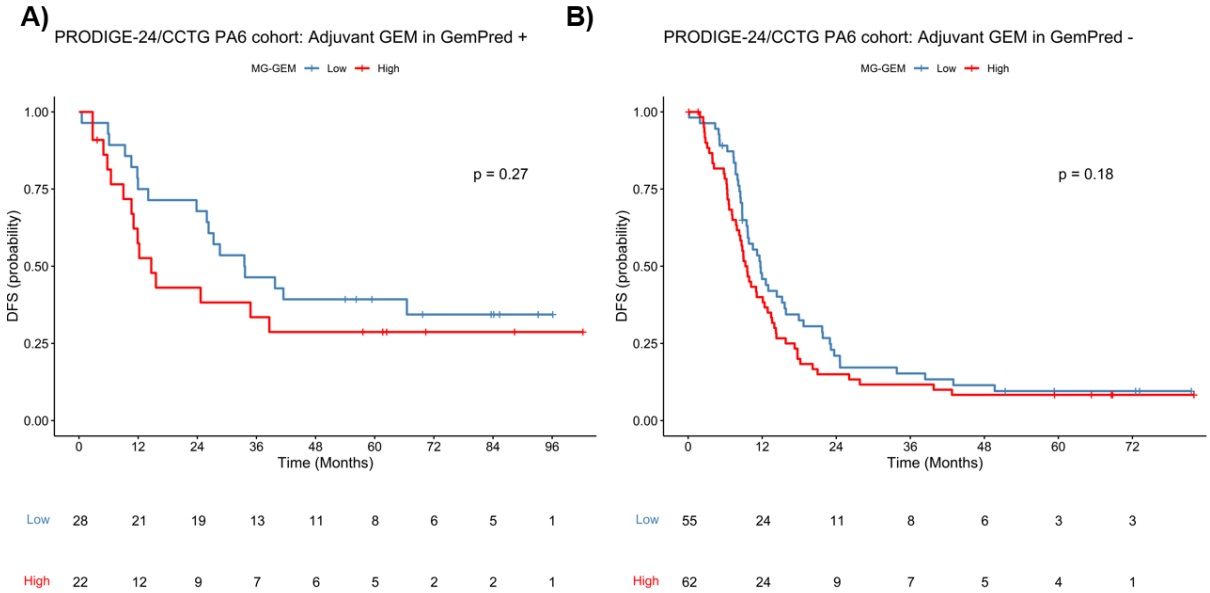


Supplementary Figure 9. OS and DFS results of the combined MG-GEM and GemPred transcriptomic signatures

(A, B) Kaplan-Meier analysis of patients without adjuvant treatment in the Puleo validation cohort showing no significant differences in overall survival (OS, A) or disease-free survival (DFS, B) between high and low MG-GEM groups.

Abbreviations: MG, Methylglyoxal; DFS, Disease-Free Survival; AT, Adjuvant Therapy; GEM, Gemcitabine.

Supp Figure 10



Supplementary Figure 10. DFS results of the combined MG-GEM and GemPred transcriptomic signatures

(A, B) Kaplan-Meier analysis of the adjuvant gemcitabine–treated PRODIGE-24/CCTG cohort showing no significant differences in disease-free survival (DFS) between high MG-GEM groups within the GemPred+ (A) and GemPred– (B) subgroups.

Abbreviations: DFS, Disease-Free Survival; MG, Methylglyoxal; AT, Adjuvant Therapy; GEM, Gemcitabine.

References

[1] L. Rahib, B. D. Smith, R. Aizenberg, A. B. Rosenzweig, J. M. Fleshman, and L. M. Matrisian, 'Projecting cancer incidence and deaths to 2030: The unexpected burden of thyroid, liver, and pancreas cancers in the united states', Jun. 01, 2014, *American Association for Cancer Research Inc.* doi: 10.1158/0008-5472.CAN-14-0155.

[2] M. Dalmartello *et al.*, 'European cancer mortality predictions for the year 2022 with focus on ovarian cancer', *Annals of Oncology*, vol. 33, no. 3, pp. 330–339, Mar. 2022, doi: 10.1016/j.annonc.2021.12.007.

[3] T. Conroy *et al.*, 'FOLFIRINOX or Gemcitabine as Adjuvant Therapy for Pancreatic Cancer', *New England Journal of Medicine*, vol. 379, no. 25, pp. 2395–2406, Dec. 2018, doi: 10.1056/nejmoa1809775.

[4] M. A. Tempero *et al.*, 'Adjuvant nab-Paclitaxel 1 Gemcitabine in Resected Pancreatic Ductal Adenocarcinoma: Results From a Randomized, Open-Label, Phase III Trial', *J Clin Oncol*, vol. 41, pp. 2007–2019, 2022, doi: 10.1200/JCO.22.

[5] Z. A. Wainberg *et al.*, 'NALIRIFOX versus nab-paclitaxel and gemcitabine in treatment-naive patients with metastatic pancreatic ductal adenocarcinoma (NAPOLI 3): a randomised, open-label, phase 3 trial', *The Lancet*, vol. 402, no. 10409, pp. 1272–1281, Oct. 2023, doi: 10.1016/S0140-6736(23)01366-1.

[6] S. Benkhaled *et al.*, 'Combination, Modulation and Interplay of Modern Radiotherapy with the Tumor Microenvironment and Targeted Therapies in Pancreatic Cancer: Which Candidates to Boost Radiotherapy?', Feb. 01, 2023, *MDPI*. doi: 10.3390/cancers15030768.

[7] M. Hilmi *et al.*, 'The immunological landscape in pancreatic ductal adenocarcinoma and overcoming resistance to immunotherapy', *Lancet Gastroenterol Hepatol*, vol. 8, no. 12, pp. 1129–1142, 2023, doi: https://doi.org/10.1016/S2468-1253(23)00207-8.

[8] A. Lambert *et al.*, 'An update on treatment options for pancreatic adenocarcinoma', 2019, *SAGE Publications Inc.* doi: 10.1177/1758835919875568.

[9] R. Nicolle *et al.*, 'A transcriptomic signature to predict adjuvant gemcitabine sensitivity in pancreatic adenocarcinoma', *Annals of Oncology*, vol. 32, no. 2, pp. 250–260, Feb. 2021, doi: 10.1016/j.annonc.2020.10.601.

[10] N. Fraunhoffer *et al.*, 'A Transcriptomic-Based Tool to Predict Gemcitabine Sensitivity in Advanced Pancreatic Adenocarcinoma', Mar. 01, 2023, *W.B. Saunders*. doi: 10.1053/j.gastro.2022.11.035.

[11] N. Fraunhoffer *et al.*, ‘Development and validation of AI-assisted transcriptomic signatures to personalize adjuvant chemotherapy in patients with pancreatic ductal adenocarcinoma’, *Annals of Oncology*, Jun. 2024, doi: 10.1016/j.annonc.2024.06.010.

[12] H. Ying *et al.*, ‘Oncogenic kras maintains pancreatic tumors through regulation of anabolic glucose metabolism’, *Cell*, vol. 149, no. 3, pp. 656–670, Apr. 2012, doi: 10.1016/j.cell.2012.01.058.

[13] M. V Liberti and J. W. Locasale, ‘The Warburg Effect: How Does it Benefit Cancer Cells?’, *Trends Biochem Sci*, vol. 41, no. 3, pp. 211–218, 2016, doi: https://doi.org/10.1016/j.tibs.2015.12.001.

[14] A. Bellahcène, M.-J. Nokin, V. Castronovo, and C. Schalkwijk, ‘Methylglyoxal-derived stress: An emerging biological factor involved in the onset and progression of cancer’, *Semin Cancer Biol*, vol. 49, pp. 64–74, 2018, doi: https://doi.org/10.1016/j.semcancer.2017.05.010.

[15] J. M. Karasinska *et al.*, ‘Altered gene expression along the glycolysis–cholesterol synthesis axis is associated with outcome in pancreatic cancer’, *Clinical Cancer Research*, vol. 26, no. 1, pp. 135–146, Jan. 2020, doi: 10.1158/1078-0432.CCR-19-1543.

[16] A. Daemen *et al.*, ‘Metabolite profiling stratifies pancreatic ductal adenocarcinomas into subtypes with distinct sensitivities to metabolic inhibitors’, *Proc Natl Acad Sci U S A*, vol. 112, no. 32, pp. E4410–E4417, Aug. 2015, doi: 10.1073/pnas.1501605112.

[17] Y. Rong *et al.*, ‘Lactate dehydrogenase A is overexpressed in pancreatic cancer and promotes the growth of pancreatic cancer cells’, *Tumor Biology*, vol. 34, no. 3, pp. 1523–1530, 2013, doi: 10.1007/s13277-013-0679-1.

[18] S. K. Shukla *et al.*, ‘MUC1 and HIF-1alpha Signaling Crosstalk Induces Anabolic Glucose Metabolism to Impart Gemcitabine Resistance to Pancreatic Cancer’, *Cancer Cell*, vol. 32, no. 1, pp. 71-87.e7, Jul. 2017, doi: 10.1016/j.ccell.2017.06.004.

[19] S. A. PHILLIPS and P. J. THORNALLEY, ‘The formation of methylglyoxal from triose phosphates’, *Eur J Biochem*, vol. 212, no. 1, pp. 101–105, Feb. 1993, doi: https://doi.org/10.1111/j.1432-1033.1993.tb17638.x.

[20] R. Crake *et al.*, ‘Resistance to Gemcitabine in Pancreatic Cancer Is Connected to Methylglyoxal Stress and Heat Shock Response’, *Cells*, vol. 12, no. 10, May 2023, doi: 10.3390/cells12101414.

[21] F. Puleo *et al.*, ‘Stratification of Pancreatic Ductal Adenocarcinomas Based on Tumor and Microenvironment Features’, *Gastroenterology*, vol. 155, no. 6, pp. 1999-2013.e3, Dec. 2018, doi: 10.1053/j.gastro.2018.08.033.

[22] C. Bouchart *et al.*, ‘Favorable histo-molecular remodeling of pancreatic ductal adenocarcinoma after Total Neoadjuvant Therapy including Stereotactic Body Radiotherapy’, *bioRxiv*, p. 2024.04.30.591890, Jan. 2024, doi: 10.1101/2024.04.30.591890.

[23] T. Conroy *et al.*, ‘Five-Year Outcomes of FOLFIRINOX vs Gemcitabine as Adjuvant Therapy for Pancreatic Cancer’, *JAMA Oncol*, vol. 8, no. 11, p. 1571, Nov. 2022, doi: 10.1001/jamaoncol.2022.3829.

[24] D. J. McCarthy, Y. Chen, and G. K. Smyth, ‘Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation’, *Nucleic Acids Res*, vol. 40, no. 10, pp. 4288–4297, May 2012, doi: 10.1093/nar/gks042.

[25] M. E. Ritchie *et al.*, ‘Limma powers differential expression analyses for RNA-sequencing and microarray studies’, *Nucleic Acids Res*, vol. 43, no. 7, p. e47, Jan. 2015, doi: 10.1093/nar/gkv007.

[26] R. Nicolle *et al.*, ‘Establishment of a pancreatic adenocarcinoma molecular gradient (PAMG) that predicts the clinical outcome of pancreatic cancer’, *EBioMedicine*, vol. 57, Jul. 2020, doi: 10.1016/j.ebiom.2020.102858.

[27] J. Friedman, T. Hastie, and R. Tibshirani, ‘Regularization Paths for Generalized Linear Models via Coordinate Descent’, 2010. [Online]. Available: http://www.jstatsoft.org/

[28] Y. Hao *et al.*, ‘Dictionary learning for integrative, multimodal and scalable single-cell analysis’, *Nat Biotechnol*, vol. 42, no. 2, pp. 293–304, Feb. 2024, doi: 10.1038/s41587-023-01767-y.

[29] I. Korsunsky *et al.*, ‘Fast, sensitive and accurate integration of single-cell data with Harmony’, *Nat Methods*, vol. 16, no. 12, pp. 1289–1296, Dec. 2019, doi: 10.1038/s41592-019-0619-0.

[30] A. Tourneur *et al.*, ‘Single nuclei and spatial transcriptomes suggest a stratification of papillary and anaplastic thyroid cancer cells’, *bioRxiv*, p. 2024.02.15.580495, Jan. 2024, doi: 10.1101/2024.02.15.580495.

[31] P. L. Ståhl *et al.*, ‘Visualization and analysis of gene expression in tissue sections by spatial transcriptomics’, Jul. 01, 2016, *American Association for the Advancement of Science*. doi: 10.1126/science.aaf2403.

[32] K. Wong, J. F. Navarro, L. Bergenstråhle, P. L. Ståhl, and J. Lundeberg, ‘ST Spot Detector: a web-based application for automatic spot and tissue detection

for spatial Transcriptomics image datasets', *Bioinformatics*, vol. 34, no. 11, pp. 1966–1968, Jun. 2018, doi: 10.1093/bioinformatics/bty030.

[33] F. Salmén *et al.*, 'Barcoded solid-phase RNA capture for Spatial Transcriptomics profiling in mammalian tissue sections', *Nat Protoc*, vol. 13, no. 11, pp. 2501–2534, Nov. 2018, doi: 10.1038/s41596-018-0045-2.

[34] G. Sturm *et al.*, 'Comprehensive evaluation of transcriptome-based cell-type quantification methods for immuno-oncology', in *Bioinformatics*, Oxford University Press, Jul. 2019, pp. i436–i445. doi: 10.1093/bioinformatics/btz363.

[35] E. Becht *et al.*, 'Estimating the population abundance of tissue-infiltrating immune and stromal cell populations using gene expression', *Genome Biol*, vol. 17, no. 1, Oct. 2016, doi: 10.1186/s13059-016-1070-5.

[36] W. L. Hwang *et al.*, 'Single-nucleus and spatial transcriptome profiling of pancreatic cancer identifies multicellular dynamics associated with neoadjuvant treatment', *Nat Genet*, vol. 54, no. 8, pp. 1178–1191, Aug. 2022, doi: 10.1038/s41588-022-01134-8.

[37] Y. Kieffer *et al.*, 'Single-cell analysis reveals fibroblast clusters linked to immunotherapy resistance in cancer', *Cancer Discov*, vol. 10, no. 9, pp. 1330–1351, Sep. 2020, doi: 10.1158/2159-8290.CD-19-1384.

[38] R. Nicolle *et al.*, 'A transcriptomic signature to predict adjuvant gemcitabine sensitivity in pancreatic adenocarcinoma', *Annals of Oncology*, vol. 32, no. 2, pp. 250–260, Feb. 2021, doi: 10.1016/j.annonc.2020.10.601.

[39] R. C. Grant *et al.*, 'GATA6 immunohistochemistry and prognosis after surgical resection of pancreatic adenocarcinoma: results from the ESPAC-4 trial', *ESMO Gastrointestinal Oncology*, vol. 8, Jun. 2025, doi: 10.1016/j.esmogo.2025.100138.

[40] E. A. Collisson *et al.*, 'Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy', *Nat Med*, vol. 17, no. 4, pp. 500–503, Apr. 2011, doi: 10.1038/nm.2344.

[41] P. Bailey *et al.*, 'Genomic analyses identify molecular subtypes of pancreatic cancer', *Nature*, vol. 531, no. 7592, pp. 47–52, Mar. 2016, doi: 10.1038/nature16965.

[42] R. A. Moffitt *et al.*, 'Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma', *Nat Genet*, vol. 47, no. 10, pp. 1168–1178, Sep. 2015, doi: 10.1038/ng.3398.

[43] S. Martens *et al.*, 'Different shades of pancreatic ductal adenocarcinoma, different paths towards precision therapeutic applications', *Annals of Oncology*, vol. 30, no. 9, pp. 1428–1436, Sep. 2019, doi: 10.1093/annonc/mdz181.

[44] M. Chan-Seng-Yue *et al.*, 'Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution', *Nat Genet*, vol. 52, no. 2, pp. 231–240, Feb. 2020, doi: 10.1038/s41588-019-0566-9.

[45] I. Ischenko *et al.*, 'KRAS drives immune evasion in a genetic model of pancreatic cancer', *Nat Commun*, vol. 12, no. 1, Dec. 2021, doi: 10.1038/s41467-021-21736-w.

[46] S. A. Kerk, T. Papagiannakopoulos, Y. M. Shah, and C. A. Lyssiotis, 'Metabolic networks in mutant KRAS-driven tumours: tissue specificities and the microenvironment', Aug. 01, 2021, *Nature Research*. doi: 10.1038/s41568-021-00375-9.

[47] B. Ardalan *et al.*, 'Distinct Molecular and Clinical Features of Specific Variants of KRAS Codon 12 in Pancreatic Adenocarcinoma', *Clinical Cancer Research*, vol. 31, no. 6, pp. 1082–1090, Mar. 2025, doi: 10.1158/1078-0432.CCR-24-3149.

[48] C. Liang *et al.*, 'Localisation of PGK1 determines metabolic phenotype to balance metastasis and proliferation in patients with SMAD4-negative pancreatic cancer', *Gut*, vol. 69, no. 5, p. 888, May 2020, doi: 10.1136/gutjnl-2018-317163.

[49] Y. Wang *et al.*, 'KDM5B promotes SMAD4 loss-driven drug resistance through activating DLG1/YAP to induce lipid accumulation in pancreatic ductal adenocarcinoma', *Cell Death Discov*, vol. 10, no. 1, Dec. 2024, doi: 10.1038/s41420-024-02020-4.

CORRESPONDING AUTHOR

Jean-Luc Van Laethem, Laboratory of Experimental Gastroenterology, Université Libre de Bruxelles, Route de Lennik 808, 1070 Brussels, Belgium, e-mail : jean-luc.vanlaethem@hubruxelles.be.

SUPPORT

This study was supported by a collaborative grant from the belgian National Fund for Scientific Research (FNRS) grant awarded to A.B. and J.-L.V.L. (PDR T0011.22). A.B. research is funded by grants from the FNRS, the University of Liège, and "Fondation Léon Fredericq".VD was supported by the Fondation Belge contre le cancer fundamental research grant FCC 2020-072, FNRS crédit de recherche J.0068.22 and

a ULB advanced ARC grant. C.B. research was supported by “Les Amis de l’Institut Bordet / L’Association Jules 560 Bordet” [2024-20] grant.

PRODIGE-24/CCTG PA6 was sponsored by R&D Unicancer and supported by a Clinical Research Hospital Program grant (PHRC11-006) from the French Ministry of Health and the Institut National du Cancer, and by the French National League against Cancer. . The Canadian Cancer Trials Group Pancreatic Adenocarcinoma (CCTG PA.6) part of the trial was supported by a program grant (704970) from the Canadian Cancer Society and by grants from 7 Days in May.

DATA SHARING STATEMENT

The datasets that support the findings of this study are not publicly available. Access to the dataset will be granted upon reasonable request sent to the corresponding author (OAS) and Unicancer.

AUTHOR CONTRIBUTIONS :

Conception and design:

Oier Azurmendi Senar, Akeila Bellahcène, Tatjana Arsenijevic, Jean-Luc Van-Laethem.

Financial support:

Akeila Bellahcène, Jean-Luc Van-Laethem.

Administrative support:

Marjorie Mauduit, Tatjana Arsenijevic.

Provision of study materials or patients:

Christelle Bouchart, Julie Navez, Thierry Conroy, Jérome Cros, Jean-Luc Van-Laethem.

Collection and assembly of data:

Kosta Stosic, Jawad Tarfouss, Laurine Verset, Tatjana Arsenijevic.

Data analysis and interpretation:

Oier Azurmendi Senar, Rémy Nicolle, Vincent Detours, Tatjana Arsenijevic.

Manuscript writing: All authors

Final approval of manuscript: All authors

Accountable for all aspects of the work: All authors

AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

Jerome Cros ´

Honoraria: Novartis

Consulting or Advisory Role: HISTALIM

Research Funding: ERYTECH Pharma (Inst)

Travel, Accommodations, Expenses: IPSEN, Novartis

No other potential conflicts of interest were reported.