GATA6 expression distinguishes classical and basal-like subtypes in advanced pancreatic cancer.

GATA6 expression distinguishes classical and basal-like subtypes in advanced pancreatic cancer.

Grainne M. O'Kane[1,2], Barbara T. Grünwald[1], GunHo Jang[1], Mehdi Masoomian[3], Sarah Picardo[2], Robert C. Grant[1,2], Robert E. Denroche[1], Amy Zhang[1], Yifan Wang[4,5], Bernard Lam[1], Paul Krzyzanowski[1], Illinca Lungu[1], John M.S. Bartlett[1], Melanie~~a~~ Peralta[3], Foram Vyas[3], Rama Khokha[3], James J. Biagi[6], Dianne Chadwick[7], Stephanie Ramotar[2], Shawn Hutchinson[2], Anna Dodd[2], Julie M. Wilson[1], Faiyaz Notta[1,8], George Zogopoulos[4,5], Steven Gallinger[1,2,9,10] Jennifer J. Knox[2], Sandra E. Fischer[3]

Affiliations

1. PanCuRx Translational Research Initiative, Ontario Institute for Cancer Research, Toronto, ON M5G 0A3, Canada
2. Wallace McCain Centre for Pancreatic Cancer, Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 2M9, Canada
3. Department of Pathology, University Health Network, Toronto, ON M5G 2M9, Canada
4. The Research Institute of the McGill University Health Centre, Montreal, QC H4A 3J1, Canada
5. The Goodman Cancer Research Centre of McGill University, Montreal, QC H3A 1A3, Canada
6. Kingston General Hospital, 25 King St W Kingston, ON K7L 5P9
7. Department of Laboratory Medicine and Pathobiology, University of Toronto, University Health Network, Toronto, ON M5G 2M9, Canada
8. Division of Research, Princess Margaret Cancer Centre, University Health Network, Toronto, ON M5G 2M9, Canada
9. Lunenfeld Tanenbaum Research Institute, Mount Sinai Hospital, Toronto, ON M5G 1X5, Canada
10. Hepatobiliary/Pancreatic Surgical Oncology Program, University Health Network, Toronto, ON M5G 2M9, Canada

Contact Info

Grainne M. O'Kane, MD	Grainne.O'Kane@uhn.ca
Barbara Grünwald, PhD	barbara.grunwald@uhnresearch.ca
Gun Ho Jang, PhD	GunHo.Jang@oicr.on.ca
Mehdi Masoomian, MD	mehdi.masoomian@one-mail.on.ca
Sarah Picardo,MD	sarah.picardo@uhn.ca
Robert C. Grant, MD	Robert.Grant@oicr.on.ca
Robert E. Denroche, MSc	Rob.Denroche@oicr.on.ca
Amy Zhang, MSc	Amy.Zhang@oicr.on.ca
Yifan Wang, MD	yifan.wang3@mail.mcgill.ca
Bernard Lam, PhD	Bernard.Lam@oicr.on.ca
Paul Krzyzanowski, PhD	Paul.Krzyzanowski@oicr.on.ca
Ilinca Lungu, MSc	Ilinca.lungu@oicr.on.ca
John M.S. Bartlett PhD	John.Bartlett@oicr.on.ca
Melania Peralta	Melanie.Peralta@uhn.ca
Foram Vyas	f.vyas@mail.utoronto.ca
Rama Khokha,PhD	rama.khokha@utoronto.ca
James Biagi, MD	jim.biagi@kingstonhsc.ca
Dianne Chadwick, PhD	Dianne.Chadwick@uhn.ca
Ramotar, Stephanie , MSc	Stephanie.Ramotar@uhn.ca
Hutchinson, Shawn	Shawn.Hutchinson@uhn.ca
Anna Dodd, CCRP	Anna.Dodd@uhn.ca
Julie M. Wilson PhD	Julie.wilson@oicr.on.ca
Faiyaz Notta, PhD	Faiyaz.notta@uhnresearch.ca
George Zogopoulos, MD	george.zogopoulos@mcgill.ca
Steven Gallinger, MD	steven.gallinger@uhn.ca
Jennifer J. Knox, MD	jennifer.knox@uhn.ca
Sandra E. Fischer, MD	Dr.Sandra.Fischer@uhn.ca

Corresponding Author:

Sandra E. Fischer, Laboratory Medicine Program, University Health Network, 200 Elizabeth Street, 11E207, Toronto, ON, Canada M5G 2C4. Phone: 416-340-3723; Fax: 416-340-5517; Dr.Sandra.Fischer@uhn.ca.

Acknowledgements

This study was conducted with support of the Ontario Institute for Cancer Research (PanCuRx Translational Research Initiative) through funding provided by the Government of Ontario, the Wallace McCain Centre for Pancreatic Cancer supported by the Princess Margaret Cancer Foundation, the Terry Fox Research Institute, the Canadian Cancer Society Research Institute and the Pancreatic Cancer Canada Foundation. The study was also supported by charitable donations from the Canadian Friends of the Hebrew University (Alex U. Soyka). JK is the recipient of the Lewitt Chair in Pancreatic Cancer. GOK is supported by the Lewitt fellowship. SG is the recipient of an Investigator Award from OICR. GZ is a clinical research scholar of the Fonds de recherche du Québec – Santé. BG was supported by the Princess Margaret Cancer Foundation, EMBO (ALTF 116-2018), and the Alexander-von-Humboldt Foundation (DEU 1199182 FLF-P).

- We acknowledge the contributions of team members at OICR within the Diagnostic Development platform and the Genomics & Bioinformatics platform (genomics.oicr.on.ca).

Abstract word count (250)

Manuscript word count (3300)

Abstract (250)

Purpose:

To determine the impact of basal-like and classical subtypes in advanced PDAC and to explore GATA6 expression as a surrogate biomarker.

Experimental design

Within the COMPASS trial patients proceeding to chemotherapy for advanced PDAC undergo tumour biopsy for RNA sequencing. Overall response rate (ORR) and overall survival (OS) were stratified by subtypes and according to chemotherapy received. Correlation of *GATA6* with the subtypes using gene expression profiling, in situ hybridization (ISH) were explored.

Results:

Between December 2015-May 2019, 195 patients (95%) had enough tissue for RNA sequencing; 39 (20%) were classified as basal-like and 156 (80%) as classical. RECIST response data were available for 157 patients; 29 basal-like and 128 classical where the ORR was 10% vs. 33% respectively (p=0.02). In patients with basal-like tumours treated with modified FOLFIRINOX (mFFX) (n=22) the progression rate was 60% compared to 15% in classical PDAC (p= 0.0002). Median OS in the intention to treat population (n=195) was 9.3 months for classical vs. 5.9 months for basal-like PDAC (HR 0.47 95% CI 0.32-0.69, p=0.0001). *GATA6* expression by RNAseq highly correlated with the classifier (p<0.001) and ISH predicted the subtypes with sensitivity of 89% and specificity of 83%. In a multivariable analysis, GATA6 expression was prognostic (p=0.02). In exploratory analyses, basal-like tumours, could be identified by keratin 5, were more hypoxic and enriched for a T cell inflamed gene expression signature.

Conclusions

The basal-like subtype is chemoresistant and can be distinguished from classical PDAC by GATA6 expression.

Translational relevance

The transcriptomic basal-like subtype is highly chemoresistant and patients have a shorter median overall survival compared to classical PDAC. In this study, survival was lowest in basal-like PDAC treated with modified FFX. *GATA6* expression by both RNAseq and in-situ hybridization (ISH) is highly associated with the classifier where low or absent GATA6 is seen in the basal-like subtype.

Manuscript (3300)

Introduction

By 2030, pancreatic ductal adenocarcinoma (PDAC) will become the second leading cause of cancer related mortality in North America(1). The majority of PDAC patients present with advanced disease where the mainstay of treatment remains combination chemotherapy. Modified FOLFIRINOX (FFX) and gemcitabine- nab-paclitaxel (GnP) are the most commonly used regimens resulting in median survival less than one year (2, 3). While the need to discover novel approaches is obvious, it is equally important to understand how to select the aforementioned regimens for current patients. There are no randomized data that shows superiority of either combination and patient inclusion differences are evident in the two pivotal phase III trials(2, 3). The only molecular predictor of response is prior knowledge of a pathogenic germline variant in a homologous recombination repair gene, which may influence the regimen of choice (4). Other currently targetable genomic variants are uncommon in PDAC.

Gene expression profiling, primarily in resected pancreatic tumours, describes a number of subtypes with considerable overlap, yet presently these do not inform clinical practice (5-7). Collisson et al. documented three subtypes (classical, quasimesenchymal, and exocrine-like)(6), Bailey et al. four subtypes (immunogenic, progenitor, ADEX and squamous)(5) and Moffitt et al. two subtypes (classical and basal-like)(7). The squamous (Bailey), quasimesenchymal (Collisson) and basal-like (Moffitt) cohorts align well across the classifiers and all three are associated with a poor prognosis in these studies. Despite this, varying tumour cellularity and heterogeneity in clustering methodologies leaves uncertainty as to the most appropriate classifier and furthermore, the clinical application of these subtypes to advanced stage disease is unclear.

In an effort to reconcile and apply existing knowledge, we established the COMPASS trial (Comprehensive Molecular Characterization of Advanced Pancreatic Ductal Adenocarcinomas (PDAC) for Better Treatment Selection: A Prospective Study, NCT

NCT02750657). Unique to this prospective study is the acquisition of tissue prior to chemotherapy in the advanced setting, which then undergoes laser capture microdissection (LCM) to ensure high tumour cellularity. The primary endpoint of feasibility in obtaining a high-quality genome report within 8 weeks in the first 50 patients has been published (8). In this earlier analysis, we determined that a modified Moffitt RNA signature, optimized for use in advanced stage PDAC (classical vs. basal-like, Supplementary Figure 1) may have prognostic impact (8). Furthermore, we found that *GATA6,* a transcription factor required for normal pancreas development(9), which has been shown to align with the classical subtype could represent a surrogate marker for classical PDAC (8). Here, we evaluated the modified Moffitt basal-like and classical subtypes together with *GATA6* expression on outcomes in patients receiving mFFX or GnP regimens on the expanded COMPASS trial. We further explored specific clinical and pathologic characteristics of the subtypes and evaluated *GATA6* as a surrogate biomarker and clinical tool. Post-hoc exploratory analyses were performed to seek additional positive biomarkers for the basal-like subtype.

Methods

Patient Population

The COMPASS trial is a prospective multi-institutional Canadian cohort study. Patient eligibility for the study has been previously described(8). Briefly, patients require a radiologic or histologic diagnosis of locally advanced or metastatic PDAC, suitable for combination chemotherapy, and must consent to a fresh tumor biopsy prior to treatment start. Biopsies can be taken from the primary lesion or any metastatic sites. Patients must not have had prior treatment for advanced disease. Treatment decisions are at the discretion of their medical oncologist. Response to therapy is assessed using computerized tomography (CT) and measured using RECIST 1.1. Demographics and treatment details, including subsequent treatments are prospectively collected using an electronic MEDIDATA database. This report includes all patients enrolled from December 2015 until May 2019 and follow-up censored on August 30th2019. Patients on this study were enrolled at the Princess Margaret Cancer Centre, McGill University Health Centre (MUHC) and Kingston General Hospital and the study was conducted in accordance with the Declaration of Helsinki. The COMPASS trial has been approved by participating site

Institutional Review Board (University Health Network, MUHC Centre for Applied Ethics, and Queen's University Health Sciences and Affiliated Teaching Hospitals Research Ethics Board); each patient provided written informed consent prior to study entry.

RNA sequencing and GATA6 expression

Frozen biospecimens underwent LCM for tumor enrichment. RNAseq analysis was performed at the Ontario Institute of Cancer Research as previously described (10). Briefly, reads were aligned to the human reference genome (hg38) and transcriptome (Ensembl v84) using STAR v.2.5.2a (11). Duplicated reads were marked using Picard v. 1.121 (https://github.com/broadinstitute/picard). Gene expression was calculated in fragments per kilobase of exon per million reads mapped (FPKM) using the cufflinks package v. 2.2.1 (12). A modified Moffitt classification (classical vs. basal-like) was applied to each sample with sufficient RNA for analysis (**Supplementary Figure 1**). Cut-off threshold levels for *GATA6* expression were determined using the maximal chi-squared method on RECIST response and dichotomised *GATA6* expression.

GATA6 RNA in situ hybridization (ISH)

Given our early results, the COMPASS trial was amended (01-Feb-2017) to include GATA6 staining using an RNAscope® in situ hybridization (ISH) assay (Advanced Cell Diagnostics Inc., Hayward, CA). A semi-quantitative score was used by the study pathologist (SF) (**Supplementary Figure 2A**) as previously reported (8). Scoring was applied blinded to results of the modified Moffitt classifier.

Immunohistochemical analysis of GATA6 and keratins

To provide more widely applicable diagnostic tests for PDAC subtypes, we optimized a protocol for GATA6 immunohistochemistry (IHC) (**emethods**) using a polyclonal anti-GATA6 antibody from R&D (Cat. Number AF1700), and secondary antibody from Vector (Cat. Number VECTABA5000). DAB+ (3,3-diaminobenzidine tetrahydrochloride plus, DAKO, Cat. Number K3468) was used as chromogen and nuclei were counterstained with Mayer's hematoxylin.) (**Supplementary Figure 2B**). To assess the pattern of GATA6 staining across larger tumor regions, we used whole sections (n=30) from a previously described resection cohort with matched RNAseq data (10) together with biopsies (n=41) from the advanced cohort.

In an exploratory analysis, we sought additional clinical markers to aid subtype identification; cytokeratins associated with *GATA6* expression were identified from RNAseq data and further explored by IHC (**emethods**).

Image analysis

To control for potential bias of manual scorings of both ISH or IHC, we performed image analysis on pre-annotated tumor sections using image analysis software QuPath v0.1.3 (13). Detection parameters were established on unequivocal GATA6-high versus -low versus -absent tumors and confirmed by the study pathologist. Semi-quantitative (SQ) scores were also predicted from image analysis data using the maximal chi-squared method.

Statistical Analysis

Qualitative variables were compared by Fisher's exact test, and quantitative variables by Wilcoxon rank sum test for pairwise comparison and the Kruskal-Walis test for multiple group comparison. All patients receiving at least 1 cycle of chemotherapy were included in the analysis of overall response rate (ORR). Survival curves were plotted using the Kaplan–Meier method and hazard ratios were calculated using Cox proportional hazard regressions with *p*-values calculated using the Wald statistic. All tests were two-sided. Multiple tests *p*-values were adjusted using Benjanimi and Hochberg method (14) for independent tests or Benjamini and Yekutieli method (15) for dependent tests, respectively. Statistical significance was set at $p = 0.05$. All analyses were conducted in R version 3.2 (16). Spearman correlation coefficients were ascertained for evaluating gene expression. Sensitivity, specificity and accuracy scores were computed to assess prediction quality.

Results

Patient characteristics at baseline

Between 30th December 2015 and May 30th2019, 250 patients were enrolled and 206 were eligible (**Figure 1-Consort**). Of these, 195 patients (95%) had enough tissue for RNA analysis and are included in this report. **Table 1** shows baseline clinical and pathological characteristics of those patients. Using the modified Moffitt classifier, 39 (20%) baseline tumor samples were basal-like, and 156 (80%) classical. Locally advanced disease at diagnosis was present in 24 (12%) and these cases, in this small subset, were all identified as classical (p=0.005). Liver metastases were present in 97% of basal-like tumors compared with 69% of classical tumors (which includes the locally advanced cases) (p<0.0001). Although basal-like tumors were more frequent in male patients (p=0.02) the overall sample size was small, Other characteristics were similar between the groups (**Table 1**).

Response to chemotherapy according to modified Moffitt classification.

Of the 195 patients, 14 (7%) did not receive any chemotherapy and were considered non-evaluable (NE). A further 23 patients (12%) died as a result of rapid functional decline prior to their first scan, of which 19 received only 1 cycle of chemotherapy; five of these 23 had basal-like PDAC. One patient receiving mFFX did not have measurable disease at enrolment. Accordingly, RECIST response data were available for 157 patients (81%) including 29 patients with basal-like tumours and 128 with classical tumours (**Figure 2A**). The ORR in classical PDAC was 33% vs. 10% in basal-like PDAC (p=0.02). The rates of progression by RECIST criteria at first CT image were much higher in basal-like vs. classical PDAC (52% vs. 16% <0.0001). **Figure 2A** shows the percentage change in target lesions, demonstrating chemoresistance of the basal-like subtype. In patients treated with mFFX and evaluable for response (n=91), progression was evident in 60% of basal-like vs. 15% of classical PDAC (p=0.0002) (**Figure 2B)**. The ORR was 29.6% vs. 10% in classical vs. basal-like PDAC (p=0.09). One patient in the latter group with a partial response (PR) had a germline *BRCA2* pathogenic variant and displayed genomic characteristics of homologous recombination deficiency. The numbers treated with GnP regimens and available for response were small, progression of disease was seen in 3/9 (33%) patients with basal-like vs. 8/54 (15%) with classical tumours (p=0.18) (**Figure 2C)**. The ORR was 39% vs. 11% in classical vs. basal-like PDAC respectively (p=0.14). Of note, 20/63 (32%) received additional experimental agents in this group.

Overall survival according to the modified Moffitt classification

Overall survival in the intention to treat population (n=195) is shown in **Figure 3A**. Median follow-up is 7.17 months. Median overall survival according to receipt of chemotherapy is shown in **Figure 3B.** In patients receiving mFFX (n=103) where performance status was less likely to confound results, median overall survival was 6.5 months in basal-like vs. 10.6 months in classical subgroups (HR 0.33 95% CI 0.19-0.60, p=0.0001). These observations suggest favourable impact of mFFX in classical PDAC but little impact of mFFX in the basal-like population (**Figure 3C**). In contrast, there was no difference between subgroups when treated with GnP regimens, where median overall survival, was 8.12 months in basal-like vs. 8.19 months in classical groups respectively (HR 0.80 95% CI 0.40-1.60, p=0.53) (**Figure 3D**). In a multivariable Cox proportional hazard regression analysis, the Moffitt subtype remained highly prognostic (p=0.018). Substituting GATA6 expression for the Moffit subtype also demonstrated the prognostic impact of GATA6 in the model, again supporting its use as a biomarker of the subtypes. Of note, stage (locally advanced versus metastatic) or chemotherapy type had no impact in this observational cohort study (**Supplementary Figure 3**). To further explore if there was a significant interaction between FFX or GnP and the subtypes, we performed an interaction analysis. There was no statistically significant difference to suggest one chemotherapy regimen for one particular subtype, although basal-like tumours trended toward improved survival with GnP, p=0.08. Of note, the modified Moffitt classifier used in this study, outperforms the previously published Moffitt classifier in identifying the poor prognostic basal-like subtype (**Supplementary Figure 1B**).

GATA6 expression by RNAseq and RNA ISH is associated with modified Moffitt subtypes.

GATA6 expression remained strongly associated with the modified Moffitt transcriptomic classifier (p<0.001) in this expanded cohort (**Figure 4A, left**). In addition, the proposed RNA ISH Semiquantitative (SQ) score was highly associated with GATA6 gene expression (RNASeq) (p<0.001) (**Figure 4A, right**). Matched RNAseq and ISH results were available in 106 patients (23 with basal-like, 83 with classical subtypes). SQ

scoring of *GATA6* ISH confirmed higher *GATA6* expression (74/83 score 2-4) in classical vs. basal-like PDAC (19/23 score 0-1). Furthermore, GATA6 ISH correlated with modified Moffitt with a sensitivity of 89%, specificity of 83% and accuracy of 88%. Both manual scores and modified Moffitt calls could be predicted from image analysis data with concordance of 92% and 81%, respectively, confirming reproducibility of semiquantitative assessment (n= 106). The modified Moffitt signature remained prognostic in this staining sub-cohort (**Figure 4B**); both GATA6 ISH SQ scoring (**Figure 4C**) and subtyping inferred from image analysis of GATA6 ISH (**Figure 4D**) predicted outcome in a similar manner.

GATA6 IHC may discriminate basal-like from classical PDAC

Matched IHC and ISH results were available in only 78 advanced PDAC cases. GATA6 levels by IHC and ISH were well correlated using quantitative assessment (**Supplementary Figure 4A**) and semi-quantitative scoring (concordance 88%), indicating that GATA6 protein levels mirror RNA expression and could aid subtype identification when RNA detection is not feasible. Indeed, IHC-based semi-quantitative scoring identified most patients with classical subtype tumors by strong and moderate GATA6 staining (52/63 with scores 2-4) while basal-like subtype patients mostly exhibited no or weak GATA6 staining (9/15 with scores 0-1), so that GATA6 protein detection by IHC was associated with modified Moffitt subtypes in advanced PDAC with a sensitivity of 83%, specificity of 60%, and accuracy of 78%. Once more, this was confirmed by quantitative assessment (**Supplementary Figure 4B**) and the concordance between prediction of GATA6 scoring from image analysis to manual scoring of GATA6 was 90%.

Tissue distribution of GATA6 by IHC in a subset or resectable and metastatic PDAC

Recent data are emerging that basal and classical subtypes can co-exist in PDAC (17, 18). We therefore explored potential variation in GATA6 expression patterns. We used whole sections from selected resected cases (n=30) in addition to needle biopsies (n=41). Although early stage tumours may not necessarily reflect the biology of advanced disease

adequate tumor content is available for mor complete evaluation. GATA6 staining (IHC) in resected specimens that were basal-like (n=14) and classical (n=16) also associated with the Moffitt subtypes: extensive immunopositivity for GATA6 (>50% of tumour cells with score 2 or higher) was found in 9/16 (56%) classical tumours vs. 1/14 (7%) basal-like tumours. (**Supplementary Table 1**). Interestingly, variable GATA6 immunopositivity (<50% of tumour cells with score 2 or higher) was present in 4/16 (25%) and 4/14 (28%) of classical and basal-like tumours, respectively, documenting a group where these subtypes may co-exist. This was furthermore observed in a number of advanced PDAC biopsies, which also exhibited variable GATA6 expression by ISH and IHC **(Supplementary Figure 5)**, demonstrating that regional GATA heterogeneity can exist in resectable and advanced stage tumors. These differences were also observed at the cellular level by image analysis (**Supplementary Figure 6**). In sum, GATA6 staining patterns were widely comparable across whole sections of 22/30 (73%) resection cases. Variable GATA6 immunopositivity was present in a subset of both, classical and basal-like subtypes, in resectable and advanced disease, which may point at the presence of classical and basal regions in the same tumor.

Keratin 5 may positively identify the basal-like subtype.

GATA6 positively identifies classical PDAC, but markers for the basal subtype are lacking. In an exploratory analysis, we evaluated keratin markers associated with GATA6 expression. In line with their use as basal markers in other tumor types (19, 20), keratins 15, 5/6, 23 and 14 were inversely correlated with GATA6 expression and thus the classical subtype (**Supplementary Figure 7A**). In this post-hoc analysis, none of the identified cytokeratins were superior to GATA6 in their association with modified Moffitt subgroups, including keratin 17, a prognostic marker in PDAC (21) (**Supplementary Figure 7B**). Among these, keratin 5 (CK5) demonstrated the strongest expression differences between basal-like and classical tumors and was found to be complementary to GATA6 expression in our cohort (**Supplementary 8)**. Furthermore, GATA6 and keratin 5 often demonstrated complementary staining pattern by IHC in PDAC tissues, including in 41 COMPASS biopsies and 30 resected PDAC whole sections (**Supplementary Table 1, Figure 5**). From these specimens, we observed the presence of both GATA6 and CK5 staining in a subset of cases (**Figure 5,** bottom panels). Indeed,

the intratumoral staining pattern of the two markers was predominantly inversely correlated in 149 individual tumor regions from the 30 resected cases (**Figure 6A**). Of note, this analysis revealed a small number of regions that contained considerable number of both CK5+ and GATA6+ cells (**Figure 6A**). Double immuno-staining confirmed distinct GATA6+/CK5- and GATA6-/CK5+ regions within the same tumor (**Figure 6B**) and in individual ducts (**Figure 6C**), which further support the notion that basal-like and classical programs can co-exist in the same tumor. Overall, many basal-like cases of advanced PDAC showed CK5 positivity (10/19, 53%) whereas most classical tumors (22/23, 96%) exhibited scant (<10%) or negative CK5 staining. Keratin 5 was thus highly specific and also showed remarkable intratumoral complementarity to GATA6 staining suggesting a clinically relevant biomarker of the basal-like subtype.

Additional molecular characteristics of the basal-like phenotype

Among the 195 eligible COMPASS patients, all 8 (4%) with adenosquamous histology were basal-like and stained positive for keratin 5 by IHC, with negligible *GATA6* expression by RNA ISH. We have previously shown that the basal-like subgroup is enriched in a hypoxia-associated gene signature by gene set enrichment analysis (22) and this observation is retained in this expanded dataset ($p=0.0003$). In addition, we found higher PD-L1 expression in the basal-like cohort ($p<0.001$), higher PD-1 expression ($p<0.001$) and enrichment of a T-cell inflamed signature previously reported (23, 24) ($p=0.007$) (**Supplementary Figure 9**). Tumour mutational burden was not different between groups (2.02 mutations/Mb vs 1.96 mutations/Mb) and was consistent with that seen in an unselected PDAC cohort (25).

Discussion

Combination chemotherapy is used in the treatment of most patients with advanced PDAC, yet the field is lacking robust biomarkers of outcome to guide regimen selection. Here, we show that patients with tumors of a modified 'basal-like' phenotype, or those with low *GATA6* RNA expression, have inferior outcomes compared to those with the 'classical' phenotype. The latter are accurately identified by high GATA6 expression and

positive GATA6 staining by in situ hybridization (ISH). Our data also suggests that basal-like tumours are particularly resistant to mFFX, warranting further investigation.

Both the PRODIGE4/ACCORD 11 and the MPACT PDAC trials of metastatic disease demonstrated an improved in survival with FOLFIRINOX and GnP across all sub-cohorts compared with gemcitabine alone(2, 3), yet they provide little insight into which subgroups might benefit the most. Notably, our study shows no superiority in either regimen in an unselected population with regard to survival. In the aforementioned trials, histological groups were not documented in either study, which is not unusual since many patients have a diagnosis made from very small samples or brushings. In contrast, the histological classification in resected specimens can be more easily reported and the PRODIGE24/ CCTG PA6 trial of mFFX in the adjuvant setting documented the prognostic impact of tumour grade in multivariable analysis (26) In patients receiving mFFX, those with well-differentiated tumors benefited the most (HR 0.52, 95% CI 0.34-0.81) whereas the impact in poorly differentiated tumors was not significant. Although limitations to the three-tiered histological classification (poor, moderate and well-differentiated) in PDAC have been noted (27), well-differentiated tumors highly express the classical program and *GATA6* (28).

The resistance of the basal-like subtype to mFFX is supported by a recent collaborative study by Tiriac et al (29)demonstrating that patient- derived organoid (PDO) chemotherapy signatures may predict treatment response. The signatures indicative of individual cytotoxic agents were applied to our COMPASS cohort suggesting that the basal-like cohort subgroup was most likely to have a non-oxaliplatin sensitive signature(29). We furthermore hypothesize that basal-like tumours may have limited sensitivity to 5-Fluorouracil. Martinelli et al. demonstrated *GATA6* loss in resected PDAC with a basal-like phenotype in the ESPAC-3 trial, and shorter survival in these patients when treated with adjuvant 5-Fluorouracil . This study also showed that *GATA6* low cell lines derived from patient-derived xenografts were particularly resistant to 5-FU but not gemcitabine (30). Notably oxaliplatin was not evaluated. In search of treatment alternatives, we report here that basal-like tumors had higher hypoxia scores, and higher PD-1 and PD-L1 expression with enrichment of a T cell inflamed signature (24) which may be predictive of immunotherapy response(23), suggesting a therapeutic strategy for

clinical trial design in this chemoresistant group. Similarly, triple negative breast cancers, although associated with worse outcomes, have higher levels of tumour infiltrating lymphocytes compared to hormone receptor positive, HER2-ve tumours. The impact of immune populations within subtypes in PDAC will require further investigation (31).

Clinical applicability of RNA sequencing and tumor enrichment by LCM is currently limited given tissue acquisition, cost and time to reporting. GATA6 detection from FFPE needle biopsies at diagnosis is therefore an attractive surrogate for transcriptomic classifiers. We demonstrate concordance of GATA6 ISH with the subtypes with sensitivity and specificity of over 80% in our tumor-enriched samples. Of note, the *GATA6* gene is not part of the original Moffitt subtype signatures but rather the Bailey squamous classifier, which largely overlaps with Moffitt calls in high purity samples (5). The number of tissue specimens available for matched ISH, IHC and RNAseq was low in our study (n=78, 40%). Therefore, although specificity was much lower for IHC compared to ISH, a prospective study with adequate tissue for matched analysis is needed. Recognizing that the identification of the basal-like subtype is critical and that GATA6 is a negative marker we sought additional positive keratin biomarkers that may be more feasible for the practicing clinician. Of these, keratin 5 predicted outcomes best after *GATA6* expression and was found to exhibit high complementarity to GATA6 staining pattern and RNA expression levels. Moreover, combined keratin 5 and GATA6 stainings on serial sections and by double immune-staining have consistently suggested that basal-like and classical elements can co-exist in a subset of PDAC cases, which strongly reinforces the need for a positive basal-like biomarker and has major implications for rationalizing subtype-specific treatments. We are currently evaluating combined staining of GATA6 and keratin 5 on the COMPASS trial.

Notably microdissected tissue, although impractical in laboratory medicine practice, most accurately detects tumour gene expression, with comparatively less exocrine and immune compartments compared to TCGA datasets, as recently described(32). This therefore implies that more reliable biomarkers can be determined from highly cellular specimens. CA-19.9 is the only approved biomarker for monitoring disease in the advanced setting (33) and the POLO trial has now documented a benefit for maintenance PARPi in patients with germline *BRCA* mutations(4). Robust subtyping of pancreatic

cancer will be critical to advancing the field, GATA6 as a single biomarker and highly correlated with the Moffitt classifier will now be evaluated in a prospective trial.

This study is limited by few progression biopsies to understand the stability of the subtypes under selective pressure during chemotherapy. This is especially interesting in light of the co-existence of basal-like and classical elements, documented here and elsewhere (17, 18). In addition, the numbers of basal-like tumours treated with GnP regimens is low and the GnP group is potentially confounded by performance status. The interaction term for chemotherapy type and subtype was not significant in this study although numbers were low. We therefore cannot conclude whether GnP is a better strategy in the basal-like cohort, rather our data suggests alternative therapies are urgently needed and clinical trials to evaluate this particular group are warranted. With mFFX as current treatment of choice in the adjuvant setting, understanding chemotherapy response to subtypes has increasing importance. It should also be noted that the response rates and survival between those receiving mFFX and GnP were not statistically different in this analysis. This is supported by the recent HALO trial 109-321 study where response rates and overall survival are comparable to historical outcomes with mFFX(2, 34). This furthermore supports the need to understand which populations can benefit most from these regimens and a prospective trial has now been planned.

In the major tumor types of lung and colorectal cancer, factors such as histological subtype, molecular profile and PD-L1 status can influence the choice of upfront systemic treatment in advanced disease and have resulted in survival gains (35-37). Since PDAC will soon become the second leading cause of cancer related mortality, it behooves the oncology community to invest in biomarkers helpful for selecting standard chemotherapy. In this study, we confirm the prognostic impact of the modified Moffitt subtypes and demonstrate that basal-like PDAC responds poorly to mFFX. The basal-like cohort can be accurately identified by GATA6 RNA expression, providing a putative single important biomarker in clinical trial design.

Figure Legends

Figure 1: Consort Diagram of patients enrolled and included on the COMPASS trial. 250 patients were enrolled and 232 patients underwent biopsies. Biopsy sites included liver, pancreas and peritoneum/omentum. 195 patients were eligible with RNAseq data representing the study population.

Table 1: Baseline characteristics of cases enrolled according to modified Moffitt classification (classical vs. basal-like)

Figure 2: Waterfall plots demonstrating tumour size change according to modified Moffitt classifier

A) Tumour size change in all patients included (n=194*): This includes any chemotherapy received. The Non Evaluable patients did not have imaging to determine response
B) Tumour size change in patients receiving first-line modified FOLFIRINOX (mFFX) (n=102*)
C) Tumour size change in patients receiving gemcitabine/nab-paclitaxel (GnP) regimens (n=71)

*1 patient with non-measurable disease is not included
NE: non evaluable
mFFX: modified FOLFIRINOX
GnP: gemcitabine/nab-paclitaxel
New lesions

Figure 3: Kaplan Meier overall survival curves according to modified Moffitt subtype and chemotherapy received

A) Overall survival in the intention to treat population (n=195) which includes patients who did not receive chemotherapy or who were non evaluable
B) Overall survival in patients receiving first line mFFX or GnP regimens (at least 1 cycle) and is presented according to modified Moffitt subtype (n=174). This graph integrates curves in 3C and 3D.
C) Overall survival in patients receiving ≥ 1 cycle mFFX (n=103) according to modified Moffitt subtype
D) Overall survival in patients receiving ≥ 1 cycle GnP regimens (n=71) according to modified Moffitt subtype.

Figure 4: GATA6 expression is associated with modified Moffitt subtypes in advanced PDAC

A) Gata6 expression by RNAseq versus modified Moffitt subtypes (left), and GATA6 expression by RNAseq versus GATA6 ISH (right). Scores of 0-1 reflect the basal-like subtype and 2-5 the classical subtype.
B) Kaplan Meier curve of overall survival by modified Moffitt in patients with matched tissue for RNAseq and GATA6 ISH analysis (n=106).
C) Kaplan Meier curves of overall survival by GATA6 ISH semi-quantitative analysis in patients with matched tissue for RNAseq and GATA6 ISH analysis (n=106).

D) Kaplan Meier curve of GATA6 by QuPath image analysis in those patients with matched tissue for RNAseq and GATA6 ISH (n=106).

Figure 5: Pathology images comparing GATA6 staining by ISH and IHC, together with CK5 IHC staining

A) Advanced PDAC cases:

COMP-022: Classic with glandular architecture (HE), GATA6 ISH score 3, GATA6 IHC score 2, CK5 negative (rare positive cells), magn 100x

COMPA-0234: Basal-like with squamous features (HE), GATA6 ISH score 1, GATA6 IHC score 1 (weak/focal), CK5 positive, magn 100x

COMP-0135: Basal-like with poor differentiation (HE), GATA6 ISH score 2 (variable distribution), GATA6 IHC score 2 (variable distribution), CK5 positive, magn 100x

B) Resected PDAC cases:

Expression pattern of GATA6 and CK5 in resected PDAC.

PCSI_639: Classic with glandular architecture (HE), GATA6 ISH score 3, GATA6 IHC score 2, CK5 negative, magn 100x.

PCSI_588: Basal-like with squamous features (HE), GATA6 IHC score 1 (weak/focal), CK5 positive, magn 100x.

PCSI_645: Classic with dual phenotype (glandular and squamous) on HE, GATA6 IHC score 2 (variable distribution), CK5 positive, magn 25x.

Figure 6: Tissue pattern of GATA6 and keratin 5 expression

A) IHC staining of GATA6 and keratin 5 on serial sections from resected PDAC specimen (n = 30). Representative images, magn 25x (left). Quantification of the percentage of GATA6+ or CK5+ cells, respectively, in 149 matched regions on adjacent sections (right).

B) Dual immunostaining of GATA6 (brown) or CK5 (magenta) in resected PDAC revealing distinct regions of GATA6 or CK5 immunoreactivity, magn 25x.

C) Dual immunostaining of GATA6 (brown) or CK5 (magenta) revealing GATA6 and CK5 immunoreactivity in the same tumour ducts. Resected PDAC (left); advanced PDAC (right); magn 400x.

1. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. CA Cancer J Clin. 2016;66(1):7-30.
2. Conroy T, Desseigne F, Ychou M, Bouche O, Guimbaud R, Becouarn Y, et al. FOLFIRINOX versus gemcitabine for metastatic pancreatic cancer. N Engl J Med. 2011;364(19):1817-25.
3. Von Hoff DD, Ervin T, Arena FP, Chiorean EG, Infante J, Moore M, et al. Increased survival in pancreatic cancer with nab-paclitaxel plus gemcitabine. N Engl J Med. 2013;369(18):1691-703.
4. Golan T, Hammel P, Reni M, Van Cutsem E, Macarulla T, Hall MJ, et al. Maintenance Olaparib for Germline BRCA-Mutated Metastatic Pancreatic Cancer. N Engl J Med. 2019.
5. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. Nature. 2016;531(7592):47-52.
6. Collisson EA, Sadanandam A, Olson P, Gibb WJ, Truitt M, Gu S, et al. Subtypes of pancreatic ductal adenocarcinoma and their differing responses to therapy. Nat Med. 2011;17(4):500-3.
7. Moffitt RA, Marayati R, Flate EL, Volmar KE, Loeza SG, Hoadley KA, et al. Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma. Nat Genet. 2015;47(10):1168-78.
8. Aung KL, Fischer SE, Denroche RE, Jang GH, Dodd A, Creighton S, et al. Genomics-Driven Precision Medicine for Advanced Pancreatic Cancer: Early Results from the COMPASS Trial. Clin Cancer Res. 2017.
9. Shi ZD, Lee K, Yang D, Amin S, Verma N, Li QV, et al. Genome Editing in hPSCs Reveals GATA6 Haploinsufficiency and a Genetic Interaction with GATA4 in Human Pancreatic Development. Cell Stem Cell. 2017;20(5):675-88 e6.
10. Connor AA, Denroche RE, Jang GH, Timms L, Kalimuthu SN, Selander I, et al. Association of Distinct Mutational Signatures With Correlates of Increased Immune Activity in Pancreatic Ductal Adenocarcinoma. JAMA Oncol. 2016.
11. Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, et al. STAR: ultrafast universal RNA-seq aligner. Bioinformatics. 2013;29(1):15-21.
12. Trapnell C, Williams BA, Pertea G, Mortazavi A, Kwan G, van Baren MJ, et al. Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. Nat Biotechnol. 2010;28(5):511-5.
13. Bankhead P, Loughrey MB, Fernandez JA, Dombrowski Y, McArt DG, Dunne PD, et al. QuPath: Open source software for digital pathology image analysis. Sci Rep. 2017;7(1):16878.
14. Benjamini Y, and Hochberg, Y. (1995). Controlling the false discovery rate: a practical and powerful approach to multiple testing. _Journal of the Royal Statistical Society Series B_ *57*, 289-300.
15. Benjamini Y, and Yekutieli, D. (2001). . The control of the false discovery rate in multiple testing under dependency. _Annals of Statistics_ *29*, 1165-1188.
16. R Core Team (2015). R: A language and environment for statistical computing. R Foundation for Statistical Computing V, Austria., https://www.R-project.org/. U.
17. Chan-Seng-Yue M, Kim JC, Wilson GW, Ng K, Figueroa EF, O'Kane GM, et al. Transcription phenotypes of pancreatic cancer are driven by genomic events during tumor evolution. Nat Genet. 2020.

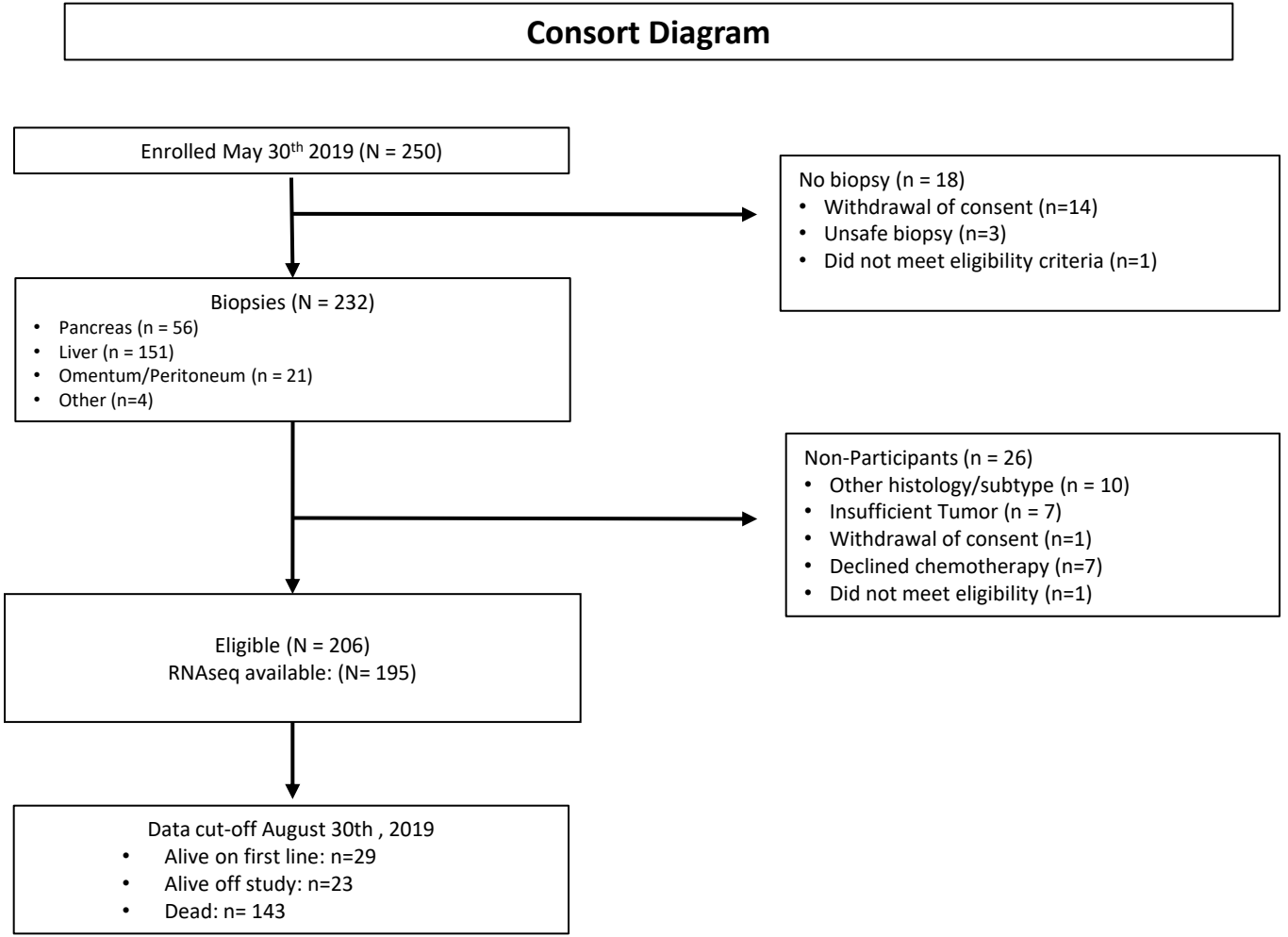
18. Hayashi A FJ, Chen R, Ho YJ, Makohon-Moore AP, Lecomte N, Zhong Y, Hong J, Huang J, Sakamoto H, Attiyeh MA. A unifying paradigm for transcriptional heterogeneity and squamous features in pancreatic ductal adenocarcinoma. Nature Cancer. 2020 Jan;1(1):59-74.
19. Abd El-Rehim DM, Pinder SE, Paish CE, Bell J, Blamey RW, Robertson JF, et al. Expression of luminal and basal cytokeratins in human breast carcinoma. J Pathol. 2004;203(2):661-71.
20. Khanom R, Sakamoto K, Pal SK, Shimada Y, Morita K, Omura K, et al. Expression of basal cell keratin 15 and keratin 19 in oral squamous neoplasms represents diverse pathophysiologies. Histol Histopathol. 2012;27(7):949-59.
21. Roa-Pena L, Leiton CV, Babu S, Pan CH, Vanner EA, Akalin A, et al. Keratin 17 identifies the most lethal molecular subtype of pancreatic cancer. Sci Rep. 2019;9(1):11239.
22. Connor AA, Denroche RE, Jang GH, Lemire M, Zhang A, Chan-Seng-Yue M, et al. Integration of Genomic and Transcriptional Features in Pancreatic Cancer Reveals Increased Cell Cycle Progression in Metastases. Cancer Cell. 2019;35(2):267-82 e7.
23. Trujillo JA, Sweis RF, Bao R, Luke JJ. T Cell-Inflamed versus Non-T Cell-Inflamed Tumors: A Conceptual Framework for Cancer Immunotherapy Drug Development and Combination Therapy Selection. Cancer Immunol Res. 2018;6(9):990-1000.
24. Spranger S, Bao R, Gajewski TF. Melanoma-intrinsic beta-catenin signalling prevents anti-tumour immunity. Nature. 2015;523(7559):231-5.
25. Cancer Genome Atlas Research Network. Electronic address aadhe, Cancer Genome Atlas Research N. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell. 2017;32(2):185-203 e13.
26. Conroy T, Hammel P, Hebbar M, Ben Abdelghani M, Wei AC, Raoul JL, et al. FOLFIRINOX or Gemcitabine as Adjuvant Therapy for Pancreatic Cancer. N Engl J Med. 2018;379(25):2395-406.
27. S NK, Wilson GW, Grant RC, Seto M, O'Kane G, Vajpeyi R, et al. Morphological classification of pancreatic ductal adenocarcinoma that predicts molecular subtypes and correlates with clinical outcome. Gut. 2019.
28. Collisson EA, Bailey P, Chang DK, Biankin AV. Molecular subtypes of pancreatic cancer. Nat Rev Gastroenterol Hepatol. 2019;16(4):207-20.
29. Tiriac H, Belleau P, Engle DD, Plenker D, Deschenes A, Somerville T, et al. Organoid profiling identifies common responders to chemotherapy in pancreatic cancer. Cancer Discov. 2018.
30. Martinelli P, Carrillo-de Santa Pau E, Cox T, Sainz B, Jr., Dusetti N, Greenhalf W, et al. GATA6 regulates EMT and tumour dissemination, and is a marker of response to adjuvant chemotherapy in pancreatic cancer. Gut. 2017;66(9):1665-76.
31. Stanton SE, Adams S, Disis ML. Variation in the Incidence and Magnitude of Tumor-Infiltrating Lymphocytes in Breast Cancer Subtypes: A Systematic Review. JAMA Oncol. 2016;2(10):1354-60.
32. Peng XL, Moffitt RA, Torphy RJ, Volmar KE, Yeh JJ. De novo compartment deconvolution and weight estimation of tumor samples using DECODER. Nat Commun. 2019;10(1):4729.
33. Winter JM, Yeo CJ, Brody JR. Diagnostic, prognostic, and predictive biomarkers in pancreatic cancer. J Surg Oncol. 2013;107(1):15-22.
34. Tempero MA VCE, Sigal D, Oh DY, Fazio N, Macarulla T, Hitre E, Hammel P, Hendifar AE, Bates SE, Li CP. HALO 109-301: A randomized, double-blind, placebo-controlled, phase 3 study of pegvorhyaluronidase alfa (PEGPH20)+ nab-

paclitaxel/gemcitabine (AG) in patients (pts) with previously untreated hyaluronan (HA)-high metastatic pancreatic ductal adenocarcinoma (mPDA). Journal of Clinical Oncology 38, no 4_suppl (February 01, 2020) 638-638.
35. Scagliotti GV, Parikh P, von Pawel J, Biesma B, Vansteenkiste J, Manegold C, et al. Phase III study comparing cisplatin plus gemcitabine with cisplatin plus pemetrexed in chemotherapy-naive patients with advanced-stage non-small-cell lung cancer. J Clin Oncol. 2008;26(21):3543-51.
36. Reck M, Rodriguez-Abreu D, Robinson AG, Hui R, Csoszi T, Fulop A, et al. Pembrolizumab versus Chemotherapy for PD-L1-Positive Non-Small-Cell Lung Cancer. N Engl J Med. 2016;375(19):1823-33.
37. Douillard JY, Siena S, Cassidy J, Tabernero J, Burkes R, Barugel M, et al. Final results from PRIME: randomized phase III study of panitumumab with FOLFOX4 for first-line treatment of metastatic colorectal cancer. Ann Oncol. 2014;25(7):1346-55.

Table 1: Baseline characteristics of patients included (n=195)

	Classical (N= 156) N, %	Basal-like (N=39) N, %	p value
Median age (yrs)	64.0 (29-84)	65.0 (44-83)	0.75
Sex			
Male	83 (53)	29 (74)	
Female	73 (47)	10 (26)	0.02
Stage			
Metastatic	132 (85)	39 (100)	
Locally advanced	24 (15)	0 (0)	0.005
Race			
White	119 (79)	27 (77)	
Asian	28 (19)	6 (17)	
African/other	4 (3)	2 (6)	
Unknown	5	4	0.23
Prior resection			
Yes	13 (8)	3 (8)	
No	143 (92)	36 (92)	0.99
CA19.9 (median, range)	1832 (1-371847)	1124 (1-71956)	0.24
Ever Smoker			
Yes	80 (51)	23 (59)	
No	76 (49)	16 (41)	0.47
Type II DM ≥18mths			
Yes	32 (21)	8 (21)	
No	120 (79)	30 (79)	0.99
Unknown	4	1	
Liver metastases			
Yes	108(69)	38 (97)	
No	48(31)	1 (3)	<0.0001
HRD genotype*			
Yes	14 (9)	2 (5)	
No	142 (91)	37 (95)	0.74
First chemotherapy			
mFolfirinox	81 (52)	22 (59)	
GnP-regimens	61 (39)	10 (26)	
Gem/nab-paclitaxel alone	*43*	*8*	
Gem/nab-paclitaxel+experimental	*18*	*2*	
Cisplatin/Gem or Gem alone	5 (3)	2 (5)	0.25
None	9 (6)	5 (13)	

Figure 1


Consort Diagram
Enrolled May 30th 2019 (N = 250)
No biopsy (n = 18)
• Withdrawal of consent (n=14)
• Unsafe biopsy (n=3)
• Did not meet eligibility criteria (n=1)
Biopsies (N = 232)
• Pancreas (n = 56)
• Liver (n = 151)
• Omentum/Peritoneum (n = 21)
• Other (n=4)
Non-Participants (n = 26)
• Other histology/subtype (n = 10)
• Insufficient Tumor (n = 7)
• Withdrawal of consent (n=1)
• Declined chemotherapy (n=7)
• Did not meet eligibility (n=1)
Eligible (N = 206)
RNAseq available: (N= 195)
Data cut-off August 30th , 2019
• Alive on first line: n=29
• Alive off study: n=23
• Dead: n= 143

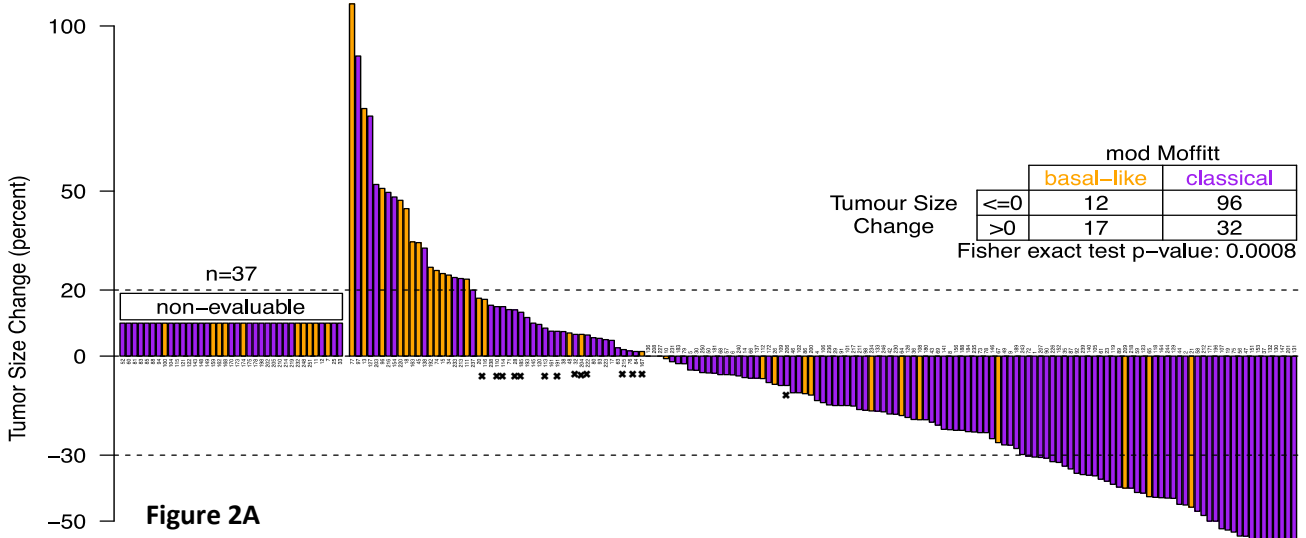


Figure 2A

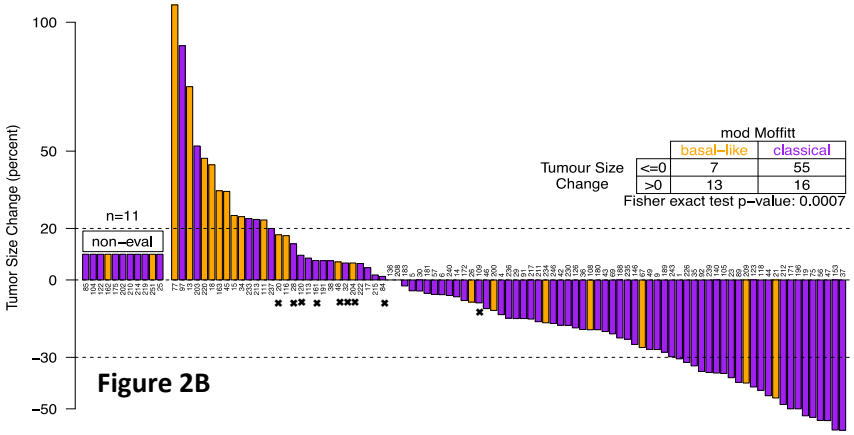


Figure 2B

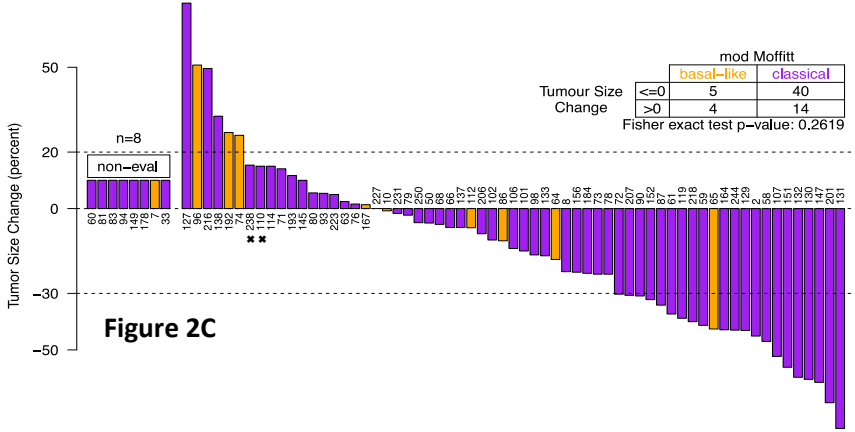


Figure 2C

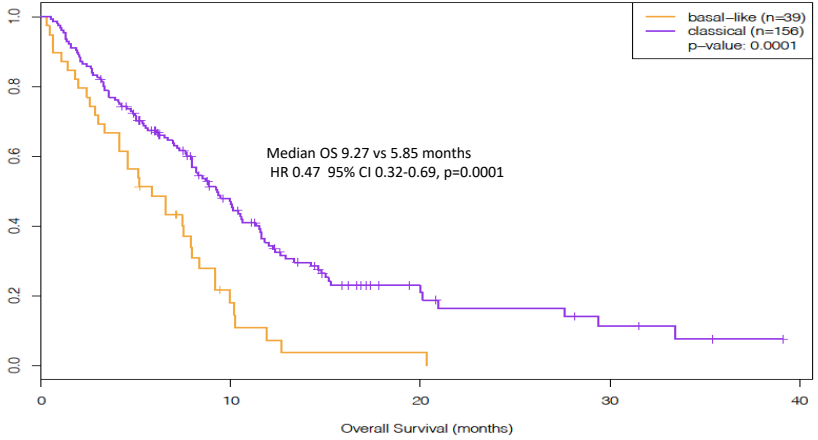


Figure 3A

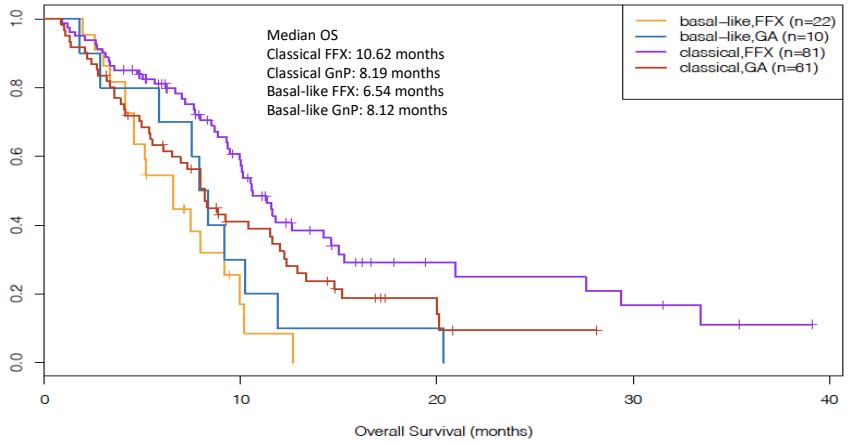


Figure 3B

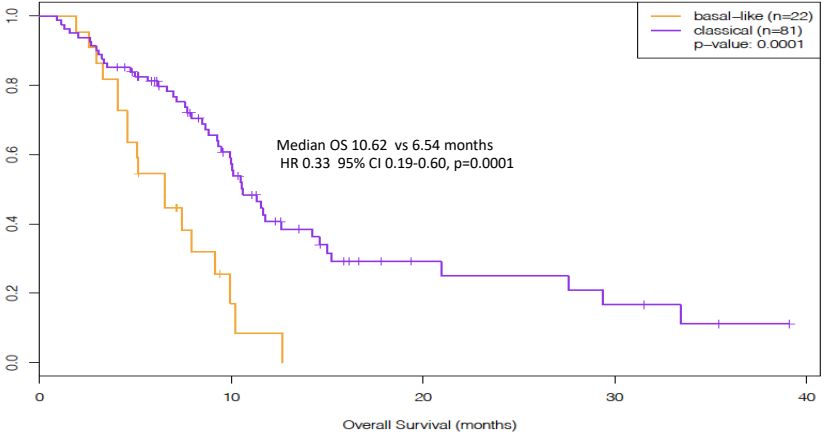


Figure 3C

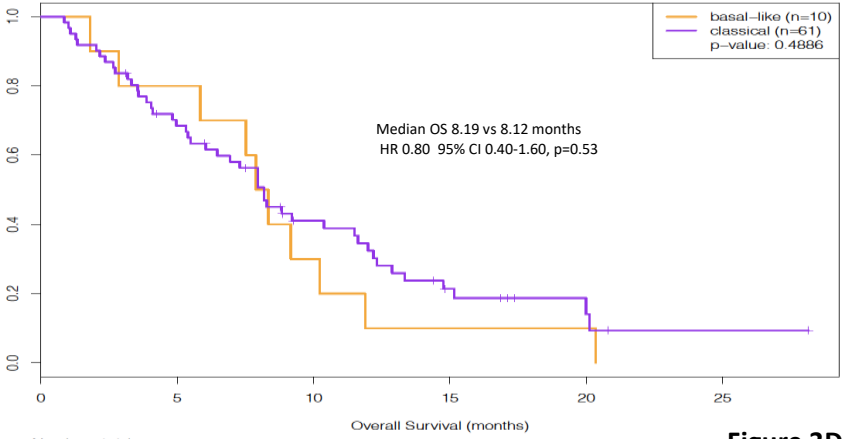
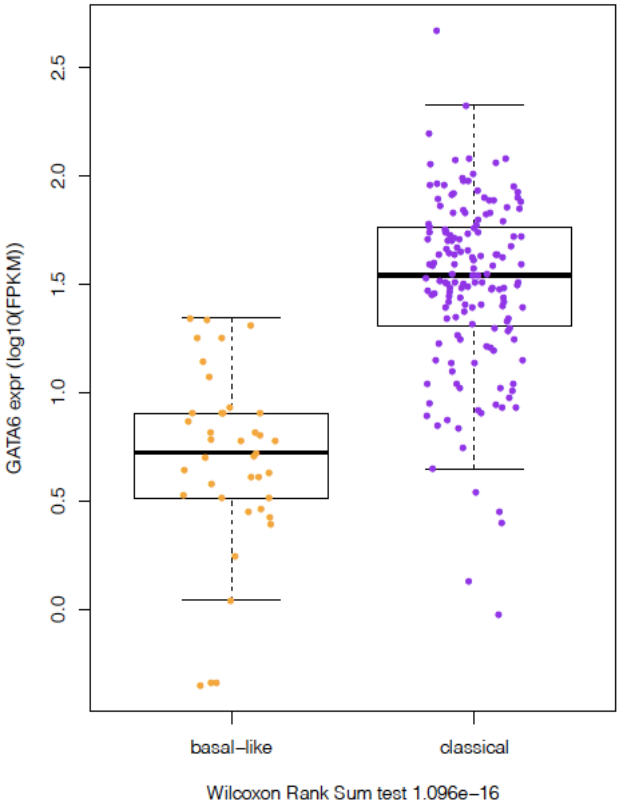


Figure 3D

Figure 4A


GATA6 expr (log10(FPKM))
0.0
0.5
1.0
1.5
2.0
2.5
basal–like
classical
Wilcoxon Rank Sum test 1.096e–16

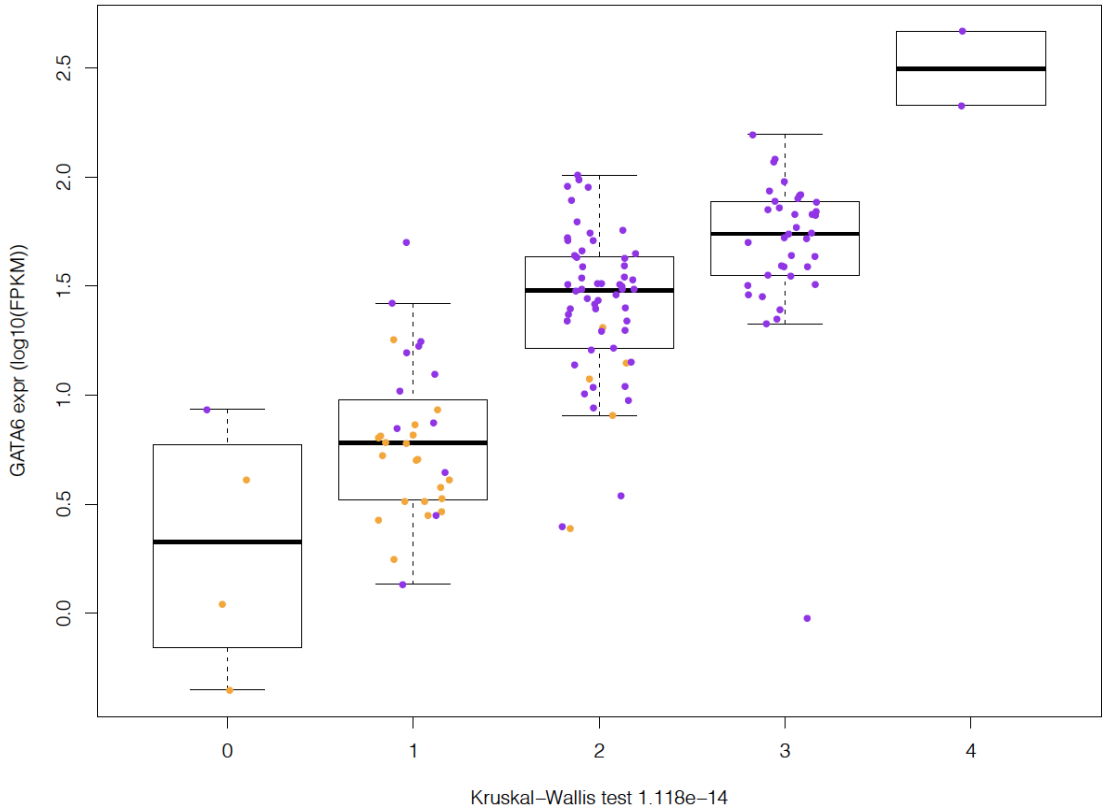

GATA6 expr (log10(FPKM))
0.0
0.5
1.0
1.5
2.0
2.5
0
1
2
3
4
Kruskal–Wallis test 1.118e–14

Figure 4B

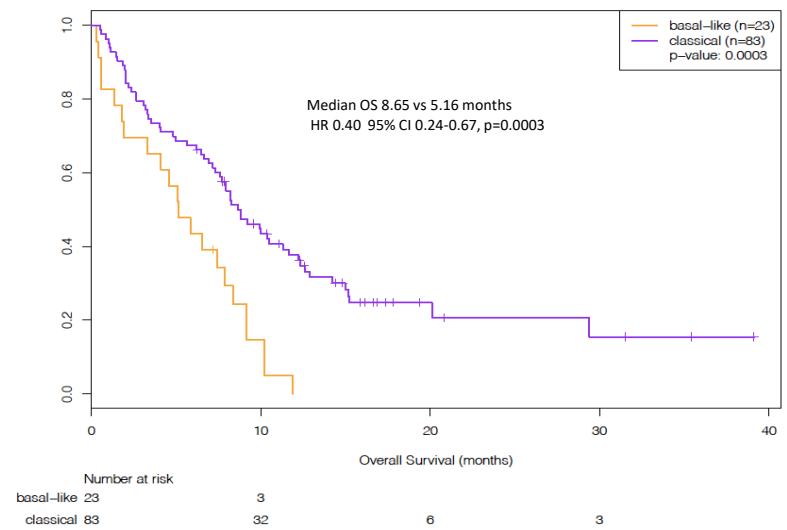


Figure 4C

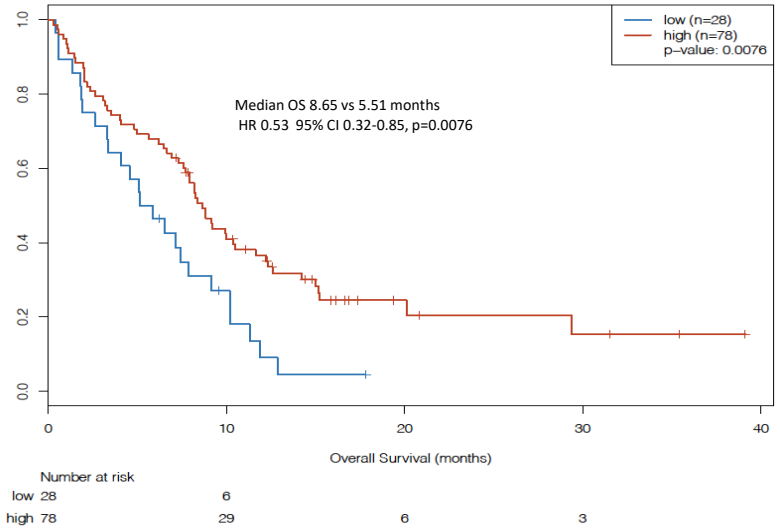


Figure 4D

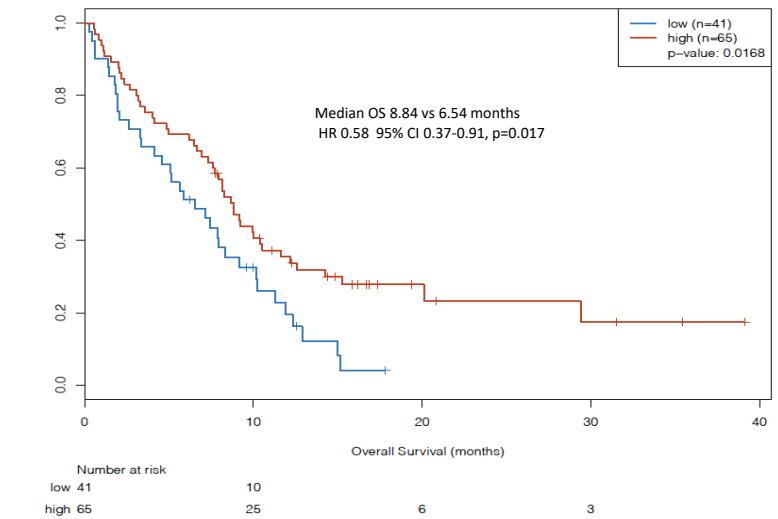


Figure 5A

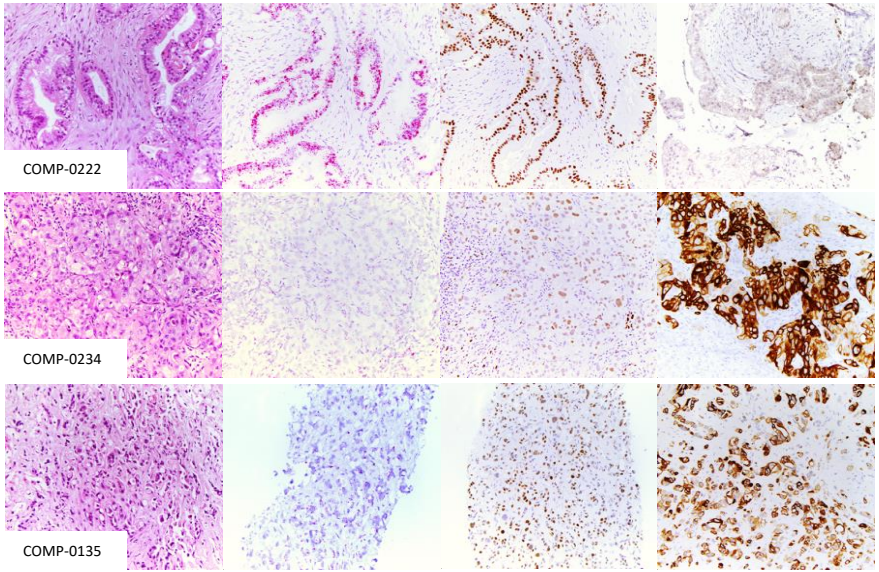


Figure 5B

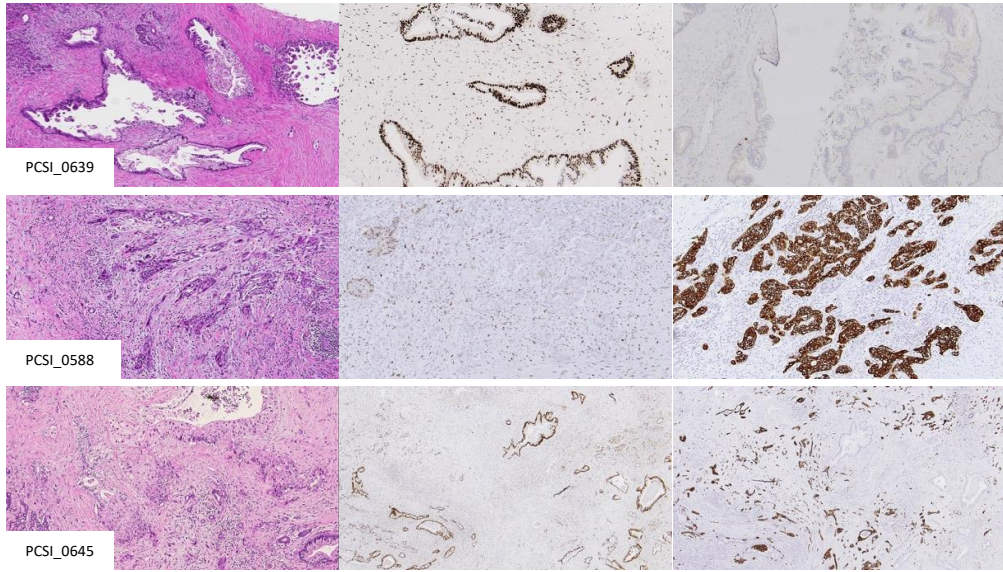
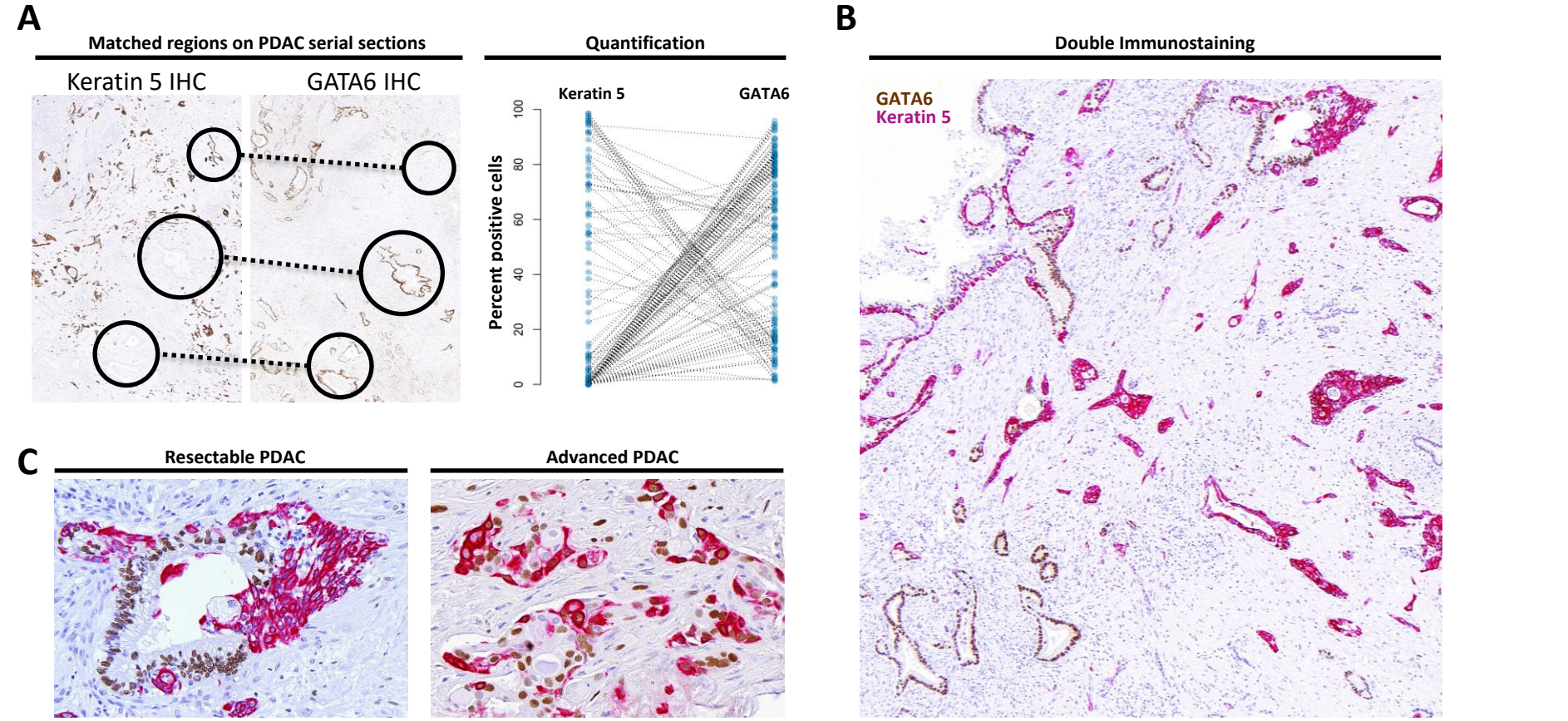


Figure 6


A
Matched regions on PDAC serial sections
Keratin 5 IHC
GATA6 IHC
Quantification
Keratin 5
GATA6
Percent positive cells
0
20
40
60
80
100
B
Double Immunostaining
GATA6
Keratin 5
C
Resectable PDAC
Advanced PDAC