

Lactate-mediated Epigenetic Reprogramming Regulates Formation of Human Pancreatic Cancer-associated Fibroblasts

Tushar D. Bhagat¹, Dagny Von Ahgrens¹, Meelad Dawlaty¹, Yiyu Zou¹, Joelle Baddour³, Abhinav Achreja³, Hongyun Zhao³, Lifeng Yang³, Brijesh Patel², Changsoo Kang⁴, Gaurav Choudhary¹, Shanisha Gordon-Mitchell¹, Srinivas Alluri¹, Sanchari Bhattacharyya¹, Srabani Sahu¹, Yiting Yu¹, Matthias Bartenstein¹, Orsi Giricz¹, , Masako Suzuki¹, Davendra Sohal⁵, Sonal Gupta⁴, Paola Guerrero⁴, Surinder Batra⁶, Michael Goggins⁷, Ulrich Steidl¹, John Greally¹, Beamon Agarwal⁸, Kith Pradhan¹, Debabrata Banerjee² , Deepak Nagrath^{9*}, Anirban Maitra^{4*}, Amit Verma^{1*}

¹ Albert Einstein College of Medicine, Montefiore Medical Center, Bronx, NY

² Rutgers University, New Brunswick, NJ

³ Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, 48109

⁴ Departments of Pathology and Translational Molecular Pathology, Sheikh Ahmed Pancreatic Cancer Research Center, UT MD Anderson Cancer Center, Houston TX

⁵ Department of Medicine, Cleveland Clinic, Cleveland, OH

⁶ University of Nebraska Medical Center, Omaha NE

⁷ Johns Hopkins, Baltimore, MD

⁸ GenomeRxUs LLC, Secane, Pennsylvania, USA

⁹ Biointerfaces Institute, University of Michigan, Ann Arbor, MI, 48109

Correspondence to:

Anirban Maitra, Departments of Pathology and Translational Molecular Pathology, Sheikh Ahmed Center for Pancreatic Cancer Research, MD Anderson Cancer Center, AMaitra@mdanderson.org

OR

Deepak Nagrath, Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI, 48109 dnagrath@umich.edu

OR

Amit Verma, Department of Developmental and Molecular Biology and Department of Medicine, Albert Einstein College of Medicine, 1300 Morris Park Avenue, Bronx, NY 10461, amit.verma@einstein.yu.edu

Key Words: Pancreatic Cancer, Stroma, Epigenetics, Lactate

Abstract (150 words)

Even though pancreatic ductal adenocarcinoma (PDAC) is associated with fibrotic stroma, the molecular pathways regulating the formation of cancer associated fibroblasts (CAFs) are not well elucidated. An epigenomic analysis of patient-derived and *de-novo* generated CAFs demonstrated widespread loss of cytosine methylation that was associated with overexpression of various inflammatory transcripts including *CXCR4*. Co-culture of neoplastic cells with CAFs led to increased invasiveness that was abrogated by inhibition of *CXCR4*. Metabolite tracing revealed that lactate produced by neoplastic cells leads to increased production of alpha-ketoglutarate (aKG) within mesenchymal stem cells (MSCs). In turn, aKG mediated activation of the demethylase TET enzyme led to decreased cytosine methylation and increased hydroxymethylation during *de novo* differentiation of MSCs to CAF. Co-injection of neoplastic cells with TET-deficient MSCs inhibited tumor growth *in vivo*. Thus, in PDAC, a tumor-mediated lactate flux is associated with widespread epigenomic reprogramming that is seen during CAF formation.

Impact Statement: Lactate in pancreatic cancer micro-environment is associated with TET enzyme activation and loss of DNA methylation seen in cancer associated fibroblasts

Introduction:

Pancreatic ductal adenocarcinoma (PDAC) is a deadly disease and is the third leading cause of deaths from cancer in the United States. An exuberant host fibrotic response, termed stromal desmoplasia, is a characteristic feature of PDAC^{1,2,3,4}. The stromal fibrosis is suspected to contribute to chemoresistance in pancreatic cancer by impeding drug delivery⁵⁻⁷, and differences in stromal behavior have been implicated with patient outcomes in pancreatic and other cancers. The stromal microenvironment predominantly consists of cancer-associated fibroblasts (CAFs) that are activated during tumorigenesis, undergoing morphological and functional changes, when compared to normal fibroblasts⁸. CAFs are derived via activation of resident pancreatic stellate cells, and also from differentiation and activation of bone marrow derived mesenchymal stem cells (BM-MSCs) that migrate to the peritumoral milieu due to chemotactic signals released by cancer cells. The characteristic features of activated CAFs include the expression of α -smooth-muscle actin (α -SMA (*ACTA1*))⁶, enhanced secretory and contractile ability, increased synthesis of extracellular matrix proteins, such as collagens, and of growth factors including basic fibroblast growth factor, transforming growth factor beta (TGF- β), interleukin-8 and platelet-derived growth factors (PDGF). In preclinical models of PDAC^{3,9,10}, as well as in other cancer types¹¹, CAFs have been shown to promote invasion and metastases through myriad mechanisms. Multiple lines of evidence support the existence of robust paracrine signals from neoplastic epithelial cells to the stromal compartment¹²⁻¹⁴, which likely facilitates the reprogramming of MSCs to an activated CAF-like state, that in turn, promotes PDAC progression. Careful studies using high density copy number arrays and mutational profiling have excluded the presence of genomic alterations in pancreatic CAFs, potentially suggesting that reprogramming is most likely epigenetic in nature¹⁵. While the epigenome of the PDAC neoplastic epithelium has been extensively studied^{16,17}, the stromal epigenome is largely uncharacterized. Thus, the underlying objective of this study was to study patterns of

epigenetic reprogramming in the PDAC stroma, specifically the most predominant α -SMA (ACTA1) expressing CAF cell type⁶.

In the present study, we studied genome wide cytosine methylation in CAFs by using both primary CAFs derived from resected PDAC, as well as *de novo* CAFs generated from MSCs *in vitro*. Our analysis revealed widespread loss of DNA methylation in CAFs as the dominant “epi-genotype”. This epigenetic reprogramming was associated with upregulation of numerous transcripts, including those encoding the chemokine receptor *CXCR4*. Our data reveal that stromal *CXCR4* overexpression promotes PDAC invasion, and provides a facile druggable target within the tumor microenvironment attenuating tumor progression. Importantly, from a mechanistic standpoint, we determine that paracrine lactate secreted by PDAC cells can be incorporated in stromal cells and lead to increased alpha-keto glutarate (aKG). This is associated with activation of the TET demethylase, thus potentially leading to epigenetic reprogramming seen during CAF formation. Our studies underscore the emerging thread between aberrant metabolism and epigenomic alterations in cancer progression, albeit from the aspect of peritumoral stroma in PDAC.

Results

Widespread epigenetic reprogramming is observed in primary and de novo transformed CAFs

Primary cultures of cancer-associated fibroblasts (CAFs) were established from 7 surgically resected PDAC tissue samples and used for epigenomic and transcriptomic analysis. Genome wide cytosine methylation was performed by the *HpaII* tiny fragment Enrichment by Ligation-mediated PCR (HELP) assay that relies on differential digestion by *HpaII* and *MspI* to identify methylated CpG sites¹⁸. Unsupervised clustering based on cytosine methylation demonstrated that pancreatic CAFs were epigenetically distinct from other non-cancer associated fibroblast controls that also included hepatic stellate cells. (**Figure 1A**). To determine the qualitative epigenetic differences between these groups we next performed a supervised analysis of the respective DNA methylation profiles. A volcano plot comparing the differences between mean methylation of individual loci between pancreatic CAFs and non-cancer associated fibroblasts demonstrated that pancreatic CAFs were characterized by widespread hypomethylation when compared to controls (5659 demethylated *versus* 674 hypermethylated loci in CAFs) (**Figure 1B**). Gene expression analyses performed on a subset of CAFs also demonstrated transcriptomic differences when compared to controls (**Figure 1C**). To elucidate the genes that were epigenetically regulated, we analyzed the genes that were concurrently overexpressed and hypomethylated in pancreatic CAFs and observed that critical cellular pathways involved in cell survival, cell cycle and cell signaling were the most significantly deregulated by epigenetically altered genes (**Supp File 1**). Multiple genes that are known to be important for cell signaling, including secreted interleukins and chemokines such as IL1a, CCL5, CCL26, cellular receptors CXCR4, ICAM3 and signaling proteins MAPK3, MAPK7, JUN were among the easily recognizable genes that exhibited differential hypomethylation and were overexpressed in pancreatic CAFs. Since striking demethylation was observed in primary CAFs, we next wanted

to validate these epigenetic changes at a higher resolution in an *in vitro* model. We generated CAFs from primary mesenchymal stem cells (MSCs) by exposing them to conditioned media from Panc-1 pancreatic cancer (PDAC-CM) cells for 21 days. This method has been shown to transform MSCs into CAFs that are functionally able to support the growth and invasion of malignant cells¹⁹ and resulted in cells with CAF like morphology and higher expression of bona fide CAF markers, α SMA (*ACTA1*) and FAP (**Figure 1D**). The methylome of MSCs and *de novo* CAFs was then studied using the HELP-tagging assay that uses massively parallel sequencing to generate genome wide CpG methylation profiles of >1.8 million loci²⁰. We observed that widespread cytosine demethylation was the dominant epigenetic change during transformation of MSCs into *de novo* CAFs (**Figure 1E,F**). Loss of methylation upon exposure to PDAC-CM was found to affect all parts of the genome (**Figure 1G,H**). Both hypo- and hypermethylated differentially methylated regions (DMRs) were distributed in various genomic locations in proportions that were comparable to the distribution of *Hpa*II sites (**Figure 1H**), thus demonstrating that epigenetic changes were occurring genome-wide during CAF transformation. This is in contrast to absence of genetic alterations in PDAC CAFs, as reported by Walter et al¹⁵.

CXCR4 is hypomethylated and overexpressed in pancreatic cancer associated fibroblasts and supports neoplastic cell invasion:

Integrative analysis between primary CAFs and *de novo* CAFs showed a common set of 130 unique promoters that were aberrantly methylated (120 hypomethylated and 10 hypermethylated) (**Figure 2A**, Venn diagram). This conserved epigenetic signature of CAFs (**Supp File 2**) was able to clearly separate these cells from normal controls and MSCs in supervised clustering (**Figure 2A**, Bottom panel). The gene promoter encoding for the CXCR4 receptor was found to be hypomethylated in both primary and *de novo* generated CAFs, and significantly demethylated CpGs were present in CpG “shores”, that have been shown to be

141 sites of differential methylation in cancer ²¹ (**Figure 2B,C**). Although recent empirical studies
142 have found upregulation of CXCR4 protein in PDAC cells by immunohistochemical assessment
143 ²², the role of this receptor on the stromal cells have not been studied.

144 To determine the functional role of CXCR4 expression on pancreatic CAFs, we used
145 specific siRNAs against CXCR4 that were able to significantly decrease *CXCR4* expression in
146 MSC-derived *de novo* CAFs (**Figure 2D, Figure 2-Figure Supplement 1**). Matrigel transwell
147 double chamber invasion assays with PDAC (PANC-1) cells revealed increased invasion of the
148 neoplastic cells in the presence of *de novo generated* CAF cells (**Figure 2E**). The increased
149 invasiveness of PDAC cells on co-culture was abrogated with RNAi-mediated knockdown of
150 *CXCR4* in the CAFs (*t*-test, *P* Value <0.05) (**Figure 2E,F, Figure 2-Figure Supplement 1**). A
151 specific inhibitor of CXCR4, AMD-3100, also led to decreased invasion of PANC-1 cells when
152 cultured with *de novo* CAFs. (**Figure 2G**). These were validated in another highly metastatic
153 Pa03C PDAC cell line (**Figure 2-Figure Supplement 1**), supporting a role for stromal CXCR4
154 blockade in attenuating tumor progression within the neoplastic compartment. Finally, gene
155 expression profiling in two sets of CAFs after *CXCR4* knockdown revealed significant alterations
156 in the transcriptional profile (**Supp File 3**), including reduced expression of IL8 and other
157 secreted ligands that are known to be tumor microenvironmental regulators ^{23,24} (**Figure 2H**).

158 **Tumor mediated lactate flux leads to production of alpha ketoglutarate, TET activation** 159 **and increased cytosine hydroxymethylation in stromal cells**

160 Next, we wanted to determine whether a diffusible factor secreted by PDAC cells could facilitate
161 the epigenomic reprogramming and demethylation of MSCs to CAFs. Lactate is produced by
162 PDAC cells via lactate dehydrogenase (LDH) enzyme during glycolysis ²⁵ and has been shown
163 to be an important mediator of metabolic pathways that can regulate the demethylase TET
164 enzymes ^{26,27}. Thus, we wanted to evaluate whether paracrine lactate, secreted by PDAC cells,

could be incorporated by MSCs and result in observed epigenetic changes through modulation of TET activity. Metabolomics analysis using uniformly ^{13}C -labeled lactate in the media as a tracer revealed that primary human MSCs can uptake lactate and convert it to pyruvate and various Krebs cycle intermediates including citrate, alpha-keto glutarate (aKG), fumarate, malate, and aspartate (**Figure 3A, Figure 3- Figure Supplement 1** showing metabolites in parallel experiments using low glucose conditions).

Since aKG is an important cofactor for TET enzymes, we next determined whether lactate secreted by PDAC-conditioned media could lead to TET activation and increased hydroxymethylcytosine (5hmC) levels in stromal cells. MSCs were exposed to conditioned media from PANC-1 cells treated with an LDH inhibitor or control media for 14 days to induce generation of *de novo* CAFs. Control PANC-1 conditioned media *per se* led to significant TET activation (**Figure 3B**) and increase in 5hmC levels in resulting *de novo* CAFs (**Figure 3C**). These effects were, however, not observed when MSCs were exposed to conditioned media from LDH inhibitor-treated PDAC cells. Conditioned media from PDAC cells with LDH knockdown with siRNAs was also able to abrogate the increase in expression of CAF markers (aSMA (*ACTA1*), FSP and Vimentin) during CAF conversion (**Supp Figure 3**). Furthermore, exposure of MSCs to exogenous lactate in media was also able to increase aKG levels, TET activity and 5hmC levels after 2 weeks of exposure (**Figure 3D,E, Figure 3- Figure Supplement 2**), demonstrating the role of this exogenous metabolite in epigenetic reprogramming of MSCs to *de novo* CAFs (**Proposed Model in Figure 3E**). TET activity increased in a dose dependant manner with exogenous lactate and was abrogated by inhibition of mitochondrial pyruvate carrier inhibitor UK5099 (**Figure 3- Figure Supplement 2**). Since 2hydroxyglutarate (2HG) and fumarate are inhibitors of TET enzymes, we also observed that exogenous lactate was able to reduce 2HG/aKG and fumarate/aKG ratios in metabolic flux

experiments (**Figure 3- Figure Supplement 1**). Additionally, exogenous cell permeable aKG was also able to increase TET activity in CAFs (**Figure 3- Figure Supplement 2**).

5hmC gains are seen during MSC to CAF conversion

To determine the genes that acquire 5hmC during CAF conversion at a high resolution, we generated CAFs from MSCs exposed to Panc1 conditioned media and used the cells for genome wide 5hmC analysis using Oxidative bisulfite sequencing (OXBS). The 5hmC gains were found to occur throughout the genome (**Figure 4A**). Genes affected by 5hmC gains were found to group into important regulatory pathways (**Figure 4B**). Gene associated with 5hmC gains included CAF markers α SMA (*ACTA1*), FSP1 (*S100A4*), and Collagen (*COL3A1*, associated with fibrotic reaction seen in pancreatic cancer) (**Figure 4C-E**). We also observed 5hmC gain at the CXCR4 promoter as well as downstream enhancer (**Figure 4F**). The important TGF- β mediator, SMAD2 also acquired 5hmC (**Figure 4G**) and a transcription factor motif analysis of all sites of 5hmC gain revealed enrichment for smad2 and smad3 binding motifs (**Table 1**). TGF-beta mediated smad activation is an important regulator of fibrosis and has been shown to be activated in CAFs ².

Increased stromal 5hmC and CXCR4 expression is seen in in primary human PDAC and murine KPC PDAC tumors

We next wanted to determine the magnitude of increased 5hmC and CXCR4 in stromal cells in large cohort of human primary PDAC samples. 5hmC (**Figure 5A**) and CXCR4 (**Figure 5B**) immunohistochemical staining was done on human PDAC TMAs and grading of intensity of stain in the tumor stromal CAFs was estimated. We observed that most CAF like cells in PDAC samples were positive for 5hmC and CXCR4 (1+ to 3+ staining intensity), (**Figure 5C**). Total PDAC samples examined for 5hmC were 254 and for CXCR4 were 261.

Next, we wanted to evaluate whether increased stromal 5hmC and CXCR4 was observed in mouse model of PDAC also. PDAC samples from KPC (*Kras* mutant, *tp53* mutant)²⁸ mouse model of were obtained and immunostained for 5hmC and CXCR4. CAF like stromal cells in PDAC tumors from KPC mice were found to be positive for both 5hmC and CXCR4 staining in all tumors examined (**Figure 5D,E**).

To validate at single cell levels, we next analyzed single cell RNA-seq (scRNAseq) data from samples obtained from precancerous low grade intraductal papillary mucinous neoplasm (LG IPMN), high grade intraductal papillary mucinous neoplasm (HG IPMN) and frank pancreatic ductal adenocarcinoma (PDAC)²⁹. Stromal cell populations positive for alphaSMA (*ACTA1*) (**Figure 5G**) and fibroblast activated protein (Figure 5H) were found to cluster distinctly from malignant cells in tSNE plots. Most of stromal cells were seen in high grade IPMN and PDAC samples (**Figure 5I**). *CXCR4* expression was seen in 14/181 (8%) stromal cells and correlated with cells with higher collagen expression, that is seen in activated CAF phenotypes⁶ (TTEst, P Val=0.02)(**Figure 5K**).

TET-deficient MSCs lead to inefficient CAF conversion and reduced tumor growth *in vivo*:

Having demonstrated that exposure to PDAC-conditioned media lead to TET activation with a concomitant increase in 5hmC levels within CAFs, we next wanted to determine the functional role of TET enzymes during MSC to *de novo* CAF differentiation. MSCs were obtained from TET2 KO mice³⁰ and controls and were co-cultured with conditioned media from murine PDAC cells derived from *Kras*, *p53* mutant tumors (“KPC” cells)³¹. WT MSCs acquired fibroblastic “CAF-like” appearance after exposure to PDAC conditioned media, while TET2 KO MSCs generally retained their original morphology (**Figure 6A**). Additionally, TET2 KO MSCs that were exposed to CM led to significantly less KPC PDAC cell invasion in matrigel when

compared to WT controls (**Figure 6B,C**). Next, to determine the functional role of TET demethylase in CAF generation *in vivo*, we co-injected murine KPC cells with MSCs from TET2 KO mice and controls into immune-deficient mice. Co-injected murine PDAC cells with TET2KO MSCs had significantly slower growth rates *in vivo* when compared to controls (**Figure 6D**). Explanted allografts with TET2 KO MSCs were significantly smaller (**Figure 6E,F**) and histologically revealed significantly less cells with a CAF phenotype, on staining with α SMA (*ACTA1*) (**Figure 6G,H**), thus demonstrating less efficient CAF conversion and tumor supporting capabilities *in vivo*.

Discussion

The tumor microenvironment plays a critical role in promoting the growth and invasion of cancer cells^{11,32}. Cancer cells can recruit MSCs from the marrow and facilitate their transformation into activated CAFs, through a plethora of paracrine signals, such as chemokines^{33,34}. One aspect of the cancer cell “secretome” within the immediate juxtatumoral milieu that has not been fully examined pertains to the role of secreted metabolites, or metabolic intermediates. We demonstrate that conversion of MSCs into CAFs is associated with widespread epigenomic reprogramming. Specifically, we establish that tumor generated lactate can be incorporated by stromal cells and can potentially induce epigenomic changes via increased production of α KG. Notably, α KG is an essential cofactor for the functionality of TET demethylases³⁵. Studies in embryonic stem cells demonstrated the ability of lactate in reprogramming stem cells via epigenetic alterations³⁵. It has also been shown that tumor lactate can cause pleiotropic effects in surrounding immune cells, as well as in the tumor cells themselves³⁶. Our data demonstrate that this diffusible factor can be a potential critical mediator that facilitates CAF differentiation in the vicinity of glycolytic proliferative tumors such as PDAC. Lactate inhibitors are being developed for anti-tumor activity^{25,37} and our data suggests that these inhibitors may act via effects on tumor microenvironment also.

Epigenetic reprogramming, which manifests as widespread loss of DNA methylation and gain of cytosine hydroxymethylation at selective promoters, is seen in both MSC-derived and primary (patient-derived) pancreatic CAFs. Increased lactic acid can result in acidic environment and it is possible that the change in pH can also influence epigenetic states. Previous studies have shown that a low acidic pH leads to increases in 2HG more than aKG³⁸. 2HG increase generally leads to decreased Tet activity. Thus in our model, a low pH (due to increase lactic acid) should not account for the high Tet activity and consequent decreased 5mC that we see; suggesting that lactic acid mediated increases in aKG and Tet activity were not influenced by changes in local pH. Loss of DNA methylation has been described mainly during developmental processes of early embryo development, and also during differentiation of hematopoietic stem cells to committed red cell progenitors^{39,40}. In fact, involvement of a large proportion of the genome by demethylation is rarely outside of developmental processes. Also, though hypomethylation has been shown in some solid tumors^{41,42}, it has not been studied at single base pair resolution in the tumor microenvironment. Epigenetic studies in the tumor microenvironment have mainly been single locus studies that have focused on hypermethylation of specific gene promoters during CAF transformation. Our findings show that widespread demethylation occurs during *in vitro* transformation of MSCs to CAFs, and is recapitulated in primary patient-derived CAF samples. In fact, genome wide analysis of 5hmC showed that 5hmC acquisition was seen in genes that have been associated with pancreatic cancer associated CAFs. 5hmC is an epigenetic modification that is obtained from oxidation of 5mC marks and is an intermediary step towards demethylation⁴³. It is postulated that 5hmC can act as an independent regulatory activating mark and is associated with sites of active transcription^{44,45}. Furthermore, our findings are consistent with a study in gastric cancer that also observed loss of methylation and was consistent with our findings.⁴⁶ A recent immunohistochemical study in a murine model of PDAC also observed loss of methylation in the microenvironment⁴⁷ though it did not study locus specific changes.

Our data shows that the chemokine receptor, CXCR4, is upregulated in CAFs and is associated with loss of promoter methylation. Interestingly the promoter CpGs that are demethylated were not located in CpG islands, but in the neighboring CpG shore. CpG shores flank CpG islands and have been shown to areas where differential methylation can occur in cancer²¹. CXCR4 is a well-studied receptor in stem cells and cancer models^{48,49} and most studies have evaluated its expression on the tumor cells. We show that in addition to its roles on tumor cells, its expression on stromal cells is also functionally important in regulating tumor cell invasion. Since host stromal cells and the cancer cells cross-talk via a large variety of soluble factors, chemokines are important mediators of these processes. It has been shown that CAFs¹⁹ as well as tumor cells³⁴ secrete SDF-1/CXCL12, which is the ligand for CXCR4. A recent study demonstrated that overexpression of the CXCR4 ligand, SDF-1 in gastric cells can induce myofibroblast expansion and is consistent with the role of CXCR4 in pancreatic cancer^{34,50}. The specific inhibitor of CXCR4, AMD-3100, is a FDA approved drug for hematopoietic stem cell mobilization. Our results demonstrate a role of this drug in inhibition of tumor stromal crosstalk. We also observed that production of secretory factors, such as interleukin-8 (IL8) by CAFs, is reduced upon CXCR4 knockdown. IL8 is a well-known regulator of cell motility and has been shown to be a regulator of pancreatic cancer cell invasion and growth^{23,51,52}.

Numerous studies have shown that depletion of stroma in PDAC can lead to enhanced efficacy of chemotherapy^{3,6,7,53}. Conversely, a murine study demonstrated that depletion of CAFs can accelerate PDAC metastases⁵⁴. As the prognosis of PDCA remains dismal, various clinical approaches are being attempted to target CAFs, and will further clarify the role of these cells in carcinogenesis. Our studies support the use of agents that disrupt the cross talk between malignant cells and the stroma and suggest that CXCR4 inhibitors may have a potential therapeutic role in pancreatic cancer that should be tested in future studies.

312

313

Methods:

DNA methylation analysis by HELP assay Genomic DNA was isolated from primary CAF samples and controls with the use of a standard high-salt procedure, and the HELP assay was carried out as previously described^{42,55}. The assay uses comparative isoschizomer profiling, interrogating cytosine methylation status on a genomic scale. Briefly, genomic DNA from the samples was digested by a methylcytosine-sensitive enzyme *HpaII* in parallel with *MspI*, which is resistant to DNA methylation, and then the *HpaII* and *MspI* products were amplified by ligation-mediated PCR. Both amplified fractions were submitted to Roche-NimbleGen, Inc. (Madison, WI) for labeling and hybridization onto a human hg17 custom-designed oligonucleotide array (50-mers) covering 25,626 *HpaII* amplifiable fragments (HAF) located at gene promoters. *HpaII* amplifiable fragments are defined as genomic sequences contained between two flanking *HpaII* sites found within 200–2,000 bp from each other. Each fragment on the array is represented by 15 individual probes distributed randomly spatially across the microarray slide. Thus the microarray covers 50,000 CpGs corresponding to 14,000 gene promoters. Signal intensities at each *HpaII* amplifiable fragment were calculated as a robust (25% trimmed) mean of their component probe-level signal intensities after intensive quality control using analytical pipelines. The $\log_2(\text{HpaII}/\text{MspI})$ was used as a representative for methylation and analyzed as a continuous variable. For most loci, each fragment was categorized as either methylated, if the centered log *HpaII*/*MspI* ratio was less than zero, or hypomethylated if on the other hand the log ratio was greater than zero.

HELP-tagging for genome-wide methylation analysis of in vitro generated CAFs: The HELP-tagging assay applies massively parallel sequencing to analyze the status of 1.8 million CpGs distributed across the entire genome^{20,56}. To perform HELP-tagging assays, DNA samples were digested with *HpaII* and ligated to customized Illumina adapters with a complementary cohesive end. These adapters also contain an *EcoP15I* site that cuts into the

adjacent sequence 27 bp away, allowing us to polish that end and ligate the other Illumina adapter for library generation by PCR. The presence of the CCGG and EcoP15I sequences at the ends of the reads allowed us to remove spurious sequences. Prior to sequencing, we performed qRT-PCR with primers that measure the proportion of adapter dimer complexes in the library, usually a very small proportion (<5%) of the total library. Following sequencing, we removed low quality or unmapped reads, piled up reads on each locus and created an output for each locus in terms of read frequency. We normalized the *HpaII* signal with that of the deeply-sequenced *MspI* profiles, as performed previously^{20,57}. Results were generated using the WASP system and linked to a local mirror of the UCSC Genome Browser for visualization. HELP-tagging data were analyzed using an automated pipeline, as described previously^{20,57}. Loci were defined in a continuous variable model, given the quantitative nature of this and comparable published assays⁵⁸. Methylation values were depicted from a range of 0 to 100, with 0 representing fully methylated to 100 representing fully hypomethylated loci.

Whole-genome hydroxymethylation analysis by oxidative bisulfite sequencing

One microgram genomic DNA from MSC and CAF cells were sonicated to 100–400 bp by Bioruptor, and 0.5% (w/w) sequencing spike-in control DNA was added thereafter and purified by Ampure XP beads. Spike-in controls were added to the adapted library. Half of the library was subjected to oxidation reaction following the manufacturer's protocol (Cambridge Epigenetix). Both oxidized and nonoxidized samples were then treated with bisulphite conversion reagent. The final PCR was performed according to the manufacturer's guide using 10 cycles of amplification, purified, and sequenced at the Einstein Epigenomics Facility. Bismark was used to map the sample reads and make methylation calls. At every base location, the 5-mC percentage was estimated by the ratio of nonconverted CpG bases to the total number of bases using the pileup of reads from the OxBS sample. For the estimation of 5-hmC, both BS

and OxBS samples were analyzed, and an estimate for the percentage of 5-hmC methylation was calculated by the difference between BS and OXBS conversions.

Since 5-hmC is a less frequent modification, for further stringency in measuring the difference in ratios, we used Fisher's exact test of proportions, using the number of converted and nonconverted reads in the BS and OxBS samples, and selected sites that have a P-value <0.05. 5-mC sites were calculated by the ratio of nonconverted bases to total bases in the OxBS sample with a biologically influenced threshold of 50%.

```
x.hmc.caf = counts.msccaf[caf.hmc > msc.hmc & caf.hmc >= 0.75 & caf.mc < msc.mc,]
```

```
x.hmc.msc = counts.msccaf[caf.hmc < msc.hmc & msc.hmc >= 0.75 & caf.mc > msc.mc,]
```

CAF hits are the loci where CAF has a higher HMC score than MSC, has an HMC score of at least .75, and has a lower MC score than MSC.

MSC hits are the loci that have a higher hmc score of at least .75, and a lower MC score.

To compare hydroxymethylation between cancer and control samples, we used Fisher's test and adjusted for multiple comparisons through the Benjamini-Hochberg procedure.

Quantitative DNA methylation analysis by MassArray Epityping. Validation of HELP microarray findings was carried out by MALDI-TOF mass spectrometry using EpiTyper by MassArray (Sequenom, CA) on bisulfite-converted DNA, as previously described^{42,59,60}. MassArray primers were designed to cover the flanking *HpaII* sites for a given locus, as well as any other *HpaII* sites found up to 2,000 bp upstream of the downstream site and up to 2,000 bp downstream of the upstream site, in order to cover all possible alternative sites of digestion.

Gene expression microarrays: Gene expression data were obtained using Affymetrix Human Genome U133A 2.0 or Plus2 GeneChips; mRNA isolation, labeling, hybridization, and quality control were carried out as described before.⁴² Raw data were processed using the Robust Multi-Averaging (RMA) algorithm and Affymetrix Expression Console software. Data are

available in the NCBI Gene Expression Omnibus database (GSE101082).

Microarray data analysis: Unsupervised clustering of HELP and gene expression data by principal component analysis was performed with the use of R 2.8.2 statistical software. Supervised analysis of the methylation data were carried out with a moderated *t*-test with Benjamini-Hochberg correction with a significance level of *P* less than .05 and an absolute difference in methylation greater than 1.5 between the means of the 2 populations (eg, MSCs vs. CAFs) to increase the likelihood of detecting biologically significant changes in methylation levels.

Gene network and gene ontology analysis: Ingenuity Pathway Analysis software and the Database for Annotation, Visualization and Integrated Discovery were used to carry out network composition analyses.

In vitro generation of CAFs: PANC-1 cells were obtained from ATCC (verified by STR authentication) and were grown in DMEM (Life technologies) + 10% heat-inactivated FBS culture medium and were mycoplasma free. Conditioned medium from pancreatic cancer cell conditioned media (CM) was harvested after 16 h and centrifuged at 3,000 rpm for 5 min and supernatant was passed through Millipore sterile 50 mL filtration system with 0.45- μ m polyvinylidene difluoride membrane. Human mesenchymal stem cells (hMSCs) were exposed to fresh pancreatic cancer cell CM continuously, with the medium changed every third day for the entire 21 day time period.

Cell lines and reagents: Primary cultures of cancer associated fibroblasts (CAFs) were established from excess tissues of surgically resected pancreatic cancers at the Johns Hopkins Hospital. The excess tissues were obtained delinked from direct patient identifiers, and primary (i.e., non-immortalized) CAFs established by passaging *in vitro*, as previously described⁶¹. All CAFs were used at early passage numbers (passages 3-6), and absence of neoplastic

epithelium was confirmed by absence of cytokeratin 19 transcripts by qRT-PCR (*data not shown*). For controls, dermal fibroblasts (Hdf), human skin fibroblasts (Hsf) and hepatic stellate (Hst) cells were obtained from ATCC. The human telomerase reverse transcriptase (hTERT) immortalized CAF line, CAF19, was a kind gift from Dr. Michael Goggins at Johns Hopkins University⁴. The human pancreatic ductal adenocarcinoma cell line Pa03C⁶², generated from a liver metastasis, was maintained in DMEM complete media supplemented with 10% FBS and 1% penicillin-streptomycin under mycoplasma free conditions. Mouse pancreatic cancer KPC cells were obtained from Dr Batra³¹. LDH inhibitor FX11 was obtained from Calbiochem and used as previously²⁵

RNA extraction and quantitative real-time reverse transcription PCR: Total RNA was extracted using RNeasy Mini Kit (Qiagen, Valencia, CA). RNA was reverse transcribed using the TaqMan One-Step RT-PCR Master Mix Reagents Kit (Applied Biosystems, Foster City, CA). Quantitative RT-PCR was carried out using a pre-designed gene expression assay for *CXCR4* (Applied Biosystems) on a StepOnePlus Real-Time PCR System. Relative fold expression was determined and normalized to *GAPDH* (Applied Biosystems) levels using the $2^{-\Delta\Delta CT}$ method⁶³.

Transfection of small interfering RNA (siRNA): *De novo* (i.e., MSC-derived) CAFs and the immortalized CAF19 cells were seeded and allowed to adhere overnight in 6-well culture plates at a density of 2×10^5 cells/well. Following overnight incubation, cells were transfected with 50nmol/L siRNA targeting *CXCR4* (Dharmacon Technologies, Thermo Fisher Scientific, Lafayette, CO) or non-targeting control siRNA (Dharmacon Technologies) using DharmaFECT 4 transfection reagent. At 24h post-transfection, cells were plated for co-culture invasion assays.

Transwell Coculture Invasion Assay: 8µm pore size inserts were coated with 100µL Matrigel (1:40 Matrigel: PBS solution) (BD Biosciences, San Jose, CA) and allowed to solidify in a

notched 24-well culture plate overnight. *De novo* CAFs or CAF19 cells were then plated in the bottom chamber at a density of 5×10^4 cells/well and allowed to adhere overnight. The media was then replaced with DMEM containing 1% FBS and Pa03C or PANC-1 cells were suspended in DMEM containing 0.5% FBS and seeded at 5×10^4 /well in the top (notched insert) chamber. Following 48hr incubation, the assay was terminated and cells migrating to the underside of the insert were fixed in ethanol and stained with 0.25% crystal violet solution. Each condition was performed in triplicate. Invasion assays were also performed in the presence of the CXCR4 antagonist, AMD3100 (Sigma Aldrich, St. Louis, MO), which was added to the lower chamber.

Analysis of global DNA methylation and hydroxymethylation by Mass

Spectrophotometer: Genomic DNA was hydrolyzed by DNA Degradase Plus (Zymo Research, CA, USA) according to the manufacturer's instructions. Digested DNA was injected onto a UPLC Zorbax Eclipse Plus C18 RRHD column (Agilent Technologies, CA). The analytes were separated by gradient elution using 5% methanol/0.1% formic acid (mobile phase A) and 100% methanol (mobile phase B) at a flow rate of 0.25 ml/min. Mobile phase B was increased from 0 to 3% in 5 minutes, to 80% in 0.5 min, kept at 80% for 2 min then switched to initial conditions in 2.5 min. The effluent from the column was directed to the Agilent 6490 Triple Quadrupole mass spectrometer (Agilent Technologies, CA). The following transitions were monitored: m/z 228.1 - >112.1 (C); m/z 242.1->126.1 (5mC) and m/z 258.1->142.1 (5hmC).

Calibration solutions with varying amounts of 5hmC (0-3%), 5mC (0-10%) and fixed amount of C, were also analyzed together with the samples. The solutions were prepared from a 200bp DNA standards containing 57 cytosines which are homogeneous for C, 5hmC or 5mC. Calibration plots of % 5hmC or %5mC vs MRM Response ratio were constructed based on the data obtained. %5hmC is obtained from the ratio of $[5hmC]/[5hmC]+[C]$. Response ratio is the

response peak area for 5hmC or 5mC divided by the combined peak areas of 5hmC, 5mC and C. The % 5hmC or 5mC in the samples were determined from the calibration plots.

TET enzymatic activity: Tet activity in nuclear cell lysates was assessed by Epigenase 5mC-Hydroxylase TET Activity/Inhibition Assay Kit (Colorimetric, Epigentek). The kit contains cofactors needed for Tet activity in vitro (Ascorbic Acid, aKG and FeNH₄SO₄) and assesses the amount of active TET enzyme in cell lysates based on efficiency of conversion of 5mC to 5hmC.

Immunohistochemistry: Tissues/cells were fixed in 10% buffered formalin, embedded in paraffin and sectioned using a microtome and mounted onto glass slides. The slides were incubated at 60 C for an hour to melt the paraffin, followed by dehydrating them through gradients of ethanol (70, 80, 90 and 100%) and 100% xylene. The samples were then treated with antigen unmasking solution (Dako pharma) followed by permeabilization with 0.3% H₂O₂ and blocked using blocking buffer (5% donkey serum and 2% BSA). Samples were then incubated overnight in the primary antibody prepared in the blocking buffer followed an incubation with appropriate HRP conjugated secondary antibody. Color development was achieved by treating the samples with diaminobenzidine (DAB) and counterstaining performed using harris hematoxyline (Dako pharma). The samples were then passaged through alcohol grades and xylene to dehydrate them, mounted using permount solution (fisher scientific) and allowed to dry overnight before the image analysis.

In vivo experiments with TET KO mouse stroma coinjection: TET2 Knockout (KO) and wild type (WT) C57B6 mice were euthanized according to protocol and their femur bones were harvested. The bone marrow was flushed and the resulting cells were grown to 30-40% cell density. Once the cells were adhered to the culture flask surface, the WT and KO stroma cells were divided into two groups. One group was treated with media conditioned with the KPC cells and the control group was treated with plain culture medium every other day for two weeks. The

resulting cells were used for downstream experiments. KPC cells (5 million/mouse) were injected along with TET2 KO stromal or WT stromal cells (1 million/mouse) into NOD-scid IL2Rgamma null (NSG) immuno-deficient mice and then followed for tumor measurements. Mice were sacrificed at end of experiment and tumors used for immunohistochemistry.

Stable metabolite isotope analysis using GC-MS:

Metabolic extraction MSCs were cultured with 10mM $^{13}\text{C}_3$ lactate (Cambridge Isotope Labs) for 48 hr. Spent medium was removed, and cells were washed with ice-cold PBS. Cells were then quenched with 400 μl methanol and 400 μl water containing 1 μg norvaline. Cells were scraped, washed with 800 μl ice-cold choloform, vortexed at 4°C for 30 min and centrifuged at 7,300 rpm for 10 min at 4°C. The aqueous portion was then collected and stored at -80°C until further analysis.

Derivatization Samples were first dried and dissolved in 30 μl of 2% methoxyamine hydrochloride in pyridine (Pierce) prior to sonication for 15min. Samples were then kept at 37°C for 2 hr, and transferred to 55°C for 1 hr following the addition of 45 μl MBTSTFA+1% TBDMCS (Pierce).

GC-MS measurements Analysis was done using an Agilent 6890 GC equipped with a 30-m Rtx-5 capillary column connected to an Agilent 5975B MS operating under electron impact ionization at 70 eV. Samples were injected at 1 μl and 270°C in splitless mode, and helium was used as the carrier gas at 1 ml/min. The heating cycle for the GC oven was as follows: 100°C for 3 min, followed by 300°C at 5°C/min temperature increase, for a total run time of 48 min per sample. Data was acquired in scan mode and the integrated signal of all potentially labeled ions was normalized by the norvaline signal and used to calculate the abundance of relative metabolites.

The mass isotopomer distribution was obtained by dividing the signal of each isotopomer by the sum of all isotopomer signals and corrected for natural abundance.

Nuclear protein extraction and In vitro TET enzymatic activity analysis

Cells were treated with various conditions and nuclear protein was then isolated from cells using the EpiQuik nuclear extraction kit (Epigentek Group Inc), according to the manufacturer's instructions. TET enzymatic activity was measured by using the ELISA-based Epigenase 5mC Hydroxylase TET Activity/Inhibition Assay Kit (Fluorometric) according to manufacturer's instructions. This technique relies on the conversion of methylated products at the bottom of the wells to hydroxymethylated products by the TET enzyme present in the nuclear extract. Thus the amount of hydroxymethylated products formed is a measure of the TET activity of the nuclear extract harvested from the cells being tested

Acknowledgements: TB was supported by The Einstein Training Program in Stem Cell Research from the Empire State Stem Cell Fund through New York State Department of Health Contract C34874GG. DVA was supported by a CA200561 T32 training grant.

References

1. Feig, C., *et al.* Targeting CXCL12 from FAP-expressing carcinoma-associated fibroblasts synergizes with anti-PD-L1 immunotherapy in pancreatic cancer. *Proc Natl Acad Sci U S A* **110**, 20212-20217 (2013).
2. Biffi, G., *et al.* IL1-Induced JAK/STAT Signaling Is Antagonized by TGFbeta to Shape CAF Heterogeneity in Pancreatic Ductal Adenocarcinoma. *Cancer Discov* **9**, 282-301 (2019).
3. Feig, C., *et al.* The pancreas cancer microenvironment. *Clin Cancer Res* **18**, 4266-4276 (2012).
4. Yu, J., *et al.* Unlike pancreatic cancer cells pancreatic cancer associated fibroblasts display minimal gene induction after 5-aza-2'-deoxycytidine. *PLoS One* **7**, e43456 (2012).
5. Provenzano, P.P., *et al.* Enzymatic targeting of the stroma ablates physical barriers to treatment of pancreatic ductal adenocarcinoma. *Cancer Cell* **21**, 418-429 (2012).
6. Ohlund, D., *et al.* Distinct populations of inflammatory fibroblasts and myofibroblasts in pancreatic cancer. *J Exp Med* **214**, 579-596 (2017).
7. von Ahrens, D., Bhagat, T.D., Nagrath, D., Maitra, A. & Verma, A. The role of stromal cancer-associated fibroblasts in pancreatic cancer. *J Hematol Oncol* **10**, 76 (2017).
8. Kalluri, R. The biology and function of fibroblasts in cancer. *Nat Rev Cancer* **16**, 582-598 (2016).

- 538 9. Hwang, R.F., *et al.* Cancer-associated stromal fibroblasts promote pancreatic tumor progression.
539 *Cancer Res* **68**, 918-926 (2008).
- 540 10. Xu, Z., *et al.* Role of pancreatic stellate cells in pancreatic cancer metastasis. *Am J Pathol* **177**,
541 2585-2596 (2010).
- 542 11. Karnoub, A.E., *et al.* Mesenchymal stem cells within tumour stroma promote breast cancer
543 metastasis. *Nature* **449**, 557-563 (2007).
- 544 12. Behrens, M.E., *et al.* The reactive tumor microenvironment: MUC1 signaling directly reprograms
545 transcription of CTGF. *Oncogene* **29**, 5667-5677 (2010).
- 546 13. Bailey, J.M., Mohr, A.M. & Hollingsworth, M.A. Sonic hedgehog paracrine signaling regulates
547 metastasis and lymphangiogenesis in pancreatic cancer. *Oncogene* **28**, 3513-3525 (2009).
- 548 14. Omary, M.B., Lugea, A., Lowe, A.W. & Pandol, S.J. The pancreatic stellate cell: a star on the rise
549 in pancreatic diseases. *J Clin Invest* **117**, 50-59 (2007).
- 550 15. Walter, K., Omura, N., Hong, S.M., Griffith, M. & Goggins, M. Pancreatic cancer associated
551 fibroblasts display normal allelotypes. *Cancer Biol Ther* **7**, 882-888 (2008).
- 552 16. Sato, N. & Goggins, M. The role of epigenetic alterations in pancreatic cancer. *J Hepatobiliary*
553 *Pancreat Surg* **13**, 286-295 (2006).
- 554 17. Goggins, M. Molecular markers of early pancreatic cancer. *J Clin Oncol* **23**, 4524-4531 (2005).
- 555 18. Figureueroa, M.E., *et al.* DNA methylation signatures identify biologically distinct subtypes in
556 acute myeloid leukemia. *Cancer Cell* **17**, 13-27 (2010).
- 557 19. Mishra, P.J., *et al.* Carcinoma-associated fibroblast-like differentiation of human mesenchymal
558 stem cells. *Cancer Res* **68**, 4331-4339 (2008).
- 559 20. Suzuki, M., *et al.* Optimized design and data analysis of tag-based cytosine methylation assays.
560 *Genome Biol* **11**, R36 (2010).
- 561 21. Irizarry, R.A., *et al.* The human colon cancer methylome shows similar hypo- and
562 hypermethylation at conserved tissue-specific CpG island shores. *Nat Genet* **41**, 178-186 (2009).
- 563 22. Bachet, J.B., *et al.* Contribution of CXCR4 and SMAD4 in predicting disease progression pattern
564 and benefit from adjuvant chemotherapy in resected pancreatic adenocarcinoma. *Ann Oncol* **23**,
565 2327-2335 (2012).
- 566 23. Matsuo, Y., *et al.* CXCL8/IL-8 and CXCL12/SDF-1alpha co-operatively promote invasiveness and
567 angiogenesis in pancreatic cancer. *Int J Cancer* **124**, 853-861 (2009).
- 568 24. Fang, Y., *et al.* IL-8-Positive Tumor-Infiltrating Inflammatory Cells Are a Novel Prognostic Marker
569 in Pancreatic Ductal Adenocarcinoma Patients. *Pancreas* **45**, 671-678 (2016).
- 570 25. Le, A., *et al.* Inhibition of lactate dehydrogenase A induces oxidative stress and inhibits tumor
571 progression. *Proc Natl Acad Sci U S A* **107**, 2037-2042 (2010).
- 572 26. Intlekofer, A.M., *et al.* Hypoxia Induces Production of L-2-Hydroxyglutarate. *Cell Metab* **22**, 304-
573 311 (2015).
- 574 27. Figureueroa, M.E., *et al.* Leukemic IDH1 and IDH2 mutations result in a hypermethylation
575 phenotype, disrupt TET2 function, and impair hematopoietic differentiation. *Cancer Cell* **18**, 553-
576 567 (2010).
- 577 28. Olive, K.P., *et al.* Inhibition of Hedgehog signaling enhances delivery of chemotherapy in a
578 mouse model of pancreatic cancer. *Science* **324**, 1457-1461 (2009).
- 579 29. Bernard, V., *et al.* Single-Cell Transcriptomics of Pancreatic Cancer Precursors Demonstrates
580 Epithelial and Microenvironmental Heterogeneity as an Early Event in Neoplastic Progression.
581 *Clin Cancer Res* **25**, 2194-2205 (2019).
- 582 30. Dawlaty, M.M., *et al.* Loss of Tet enzymes compromises proper differentiation of embryonic
583 stem cells. *Dev Cell* **29**, 102-111 (2014).

- 584 31. Torres, M.P., *et al.* Novel pancreatic cancer cell lines derived from genetically engineered mouse
585 models of spontaneous pancreatic adenocarcinoma: applications in diagnosis and therapy. *PLoS*
586 *One* **8**, e80580 (2013).
- 587 32. Orimo, A. & Weinberg, R.A. Stromal fibroblasts in cancer: a novel tumor-promoting cell type.
588 *Cell Cycle* **5**, 1597-1601 (2006).
- 589 33. Mishra, P., Banerjee, D. & Ben-Baruch, A. Chemokines at the crossroads of tumor-fibroblast
590 interactions that promote malignancy. *J Leukoc Biol* **89**, 31-39 (2011).
- 591 34. Quante, M., *et al.* Bone marrow-derived myofibroblasts contribute to the mesenchymal stem
592 cell niche and promote tumor growth. *Cancer Cell* **19**, 257-272 (2011).
- 593 35. Carey, B.W., Finley, L.W., Cross, J.R., Allis, C.D. & Thompson, C.B. Intracellular alpha-
594 ketoglutarate maintains the pluripotency of embryonic stem cells. *Nature* **518**, 413-416 (2015).
- 595 36. Matilainen, O., Quiros, P.M. & Auwerx, J. Mitochondria and Epigenetics - Crosstalk in
596 Homeostasis and Stress. *Trends Cell Biol* **27**, 453-463 (2017).
- 597 37. Rajeshkumar, N.V., *et al.* Therapeutic Targeting of the Warburg Effect in Pancreatic Cancer
598 Relies on an Absence of p53 Function. *Cancer Res* **75**, 3355-3364 (2015).
- 599 38. Nadtochiy, S.M., *et al.* Acidic pH Is a Metabolic Switch for 2-Hydroxyglutarate Generation and
600 Signaling. *J Biol Chem* **291**, 20188-20197 (2016).
- 601 39. Shearstone, J.R., *et al.* Global DNA demethylation during mouse erythropoiesis in vivo. *Science*
602 **334**, 799-802 (2011).
- 603 40. Yu, Y., *et al.* High resolution methylome analysis reveals widespread functional hypomethylation
604 during adult human erythropoiesis. *J Biol Chem* (2013).
- 605 41. Timp, W. & Feinberg, A.P. Cancer as a dysregulated epigenome allowing cellular growth
606 advantage at the expense of the host. *Nat Rev Cancer* **13**, 497-510 (2013).
- 607 42. Alvarez, H., *et al.* Widespread hypomethylation occurs early and synergizes with gene
608 amplification during esophageal carcinogenesis. *PLoS Genet* **7**, e1001356 (2011).
- 609 43. Ko, M., *et al.* Impaired hydroxylation of 5-methylcytosine in myeloid cancers with mutant TET2.
610 *Nature* **468**, 839-843 (2010).
- 611 44. Bhattacharyya, S., *et al.* Altered hydroxymethylation is seen at regulatory regions in pancreatic
612 cancer and regulates oncogenic pathways. *Genome Res* **27**, 1830-1842 (2017).
- 613 45. Madzo, J., *et al.* Hydroxymethylation at gene regulatory regions directs stem/early progenitor
614 cell commitment during erythropoiesis. *Cell Rep* **6**, 231-244 (2014).
- 615 46. Jiang, L., *et al.* Global hypomethylation of genomic DNA in cancer-associated myofibroblasts.
616 *Cancer Res* **68**, 9900-9908 (2008).
- 617 47. Shakya, R., *et al.* Hypomethylating therapy in an aggressive stroma-rich model of pancreatic
618 carcinoma. *Cancer Res* **73**, 885-896 (2013).
- 619 48. Rettig, M.P., Ansstas, G. & DiPersio, J.F. Mobilization of hematopoietic stem and progenitor cells
620 using inhibitors of CXCR4 and VLA-4. *Leukemia* **26**, 34-53 (2012).
- 621 49. Wong, D. & Korz, W. Translating an Antagonist of Chemokine Receptor CXCR4: from bench to
622 bedside. *Clin Cancer Res* **14**, 7975-7980 (2008).
- 623 50. Shibata, W., *et al.* Stromal cell-derived factor-1 overexpression induces gastric dysplasia through
624 expansion of stromal myofibroblasts and epithelial progenitors. *Gut* **62**, 192-200 (2013).
- 625 51. Delitto, D., *et al.* Human Pancreatic Cancer Cells Induce a MyD88-Dependent Stromal Response
626 to Promote a Tumor-Tolerant Immune Microenvironment. *Cancer Res* **77**, 672-683 (2017).
- 627 52. Sparmann, A. & Bar-Sagi, D. Ras-induced interleukin-8 expression plays a critical role in tumor
628 growth and angiogenesis. *Cancer Cell* **6**, 447-458 (2004).
- 629 53. Ireland, L., *et al.* Chemoresistance in Pancreatic Cancer Is Driven by Stroma-Derived Insulin-Like
630 Growth Factors. *Cancer Res* **76**, 6851-6863 (2016).

54. Ozdemir, B.C., *et al.* Depletion of carcinoma-associated fibroblasts and fibrosis induces immunosuppression and accelerates pancreas cancer with reduced survival. *Cancer Cell* **25**, 719-734 (2014).
55. Khulan, B., *et al.* Comparative isoschizomer profiling of cytosine methylation: the HELP assay. *Genome Res* **16**, 1046-1055 (2006).
56. Bhattacharyya, S., *et al.* Genome-wide hydroxymethylation tested using the HELP-GT assay shows redistribution in cancer. *Nucleic Acids Res* **41**, e157 (2013).
57. Wu, W., *et al.* Hypomethylation of noncoding DNA regions and overexpression of the long noncoding RNA, AFAP1-AS1, in Barrett's esophagus and esophageal adenocarcinoma. *Gastroenterology* **144**, 956-966 e954 (2013).
58. Ball, M.P., *et al.* Targeted and genome-scale strategies reveal gene-body methylation signatures in human cells. *Nat Biotechnol* **27**, 361-368 (2009).
59. Figureueroa, M.E., *et al.* An integrative genomic and epigenomic approach for the study of transcriptional regulation. *PLoS ONE* **3**, e1882 (2008).
60. Figureueroa, M.E., *et al.* Genome-wide epigenetic analysis delineates a biologically distinct immature acute leukemia with myeloid/T-lymphoid features. *Blood* **113**, 2795-2804 (2009).
61. Walter, K., *et al.* Overexpression of smoothened activates the sonic hedgehog signaling pathway in pancreatic cancer-associated fibroblasts. *Clin Cancer Res* **16**, 1781-1789 (2010).
62. Jones, S., *et al.* Exomic sequencing identifies PALB2 as a pancreatic cancer susceptibility gene. *Science* **324**, 217 (2009).
63. Schmittgen, T.D. & Livak, K.J. Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **3**, 1101-1108 (2008).

Figure Legends:

Figure 1: Widespread epigenetic and transcriptomic alterations are seen in pancreatic cancer associated fibroblasts:

A: Unsupervised clustering of cytosine methylation profiles from 7 primary patient-derived pancreatic cancer associated fibroblasts (CAFs) and 4 healthy controls (Hst: Hepatic stellate cells, Hdf: Human dermal fibroblasts, Hsf: Human skin fibroblasts) shows that CAFs are epigenetically distinct

B: Volcano plot shows that majority of differentially methylated loci in primary patient-derived CAFs are hypomethylated when compared to controls

C: Unsupervised clustering of gene expression profiles shows transcriptomic differences between CAFs and controls

665 D: Three independent MSCs were exposed to PANC-1 conditioned media (CM) for 21 days and
666 analyzed for α -SMA (*ACTA1*) and Fibroblast activation protein (FAP) expression. (T-test,
667 $P < 0.05$).

668 E: Unsupervised clustering based on cytosine methylation shows epigenomic differences
669 between MSCs and *de novo* generated CAFs. Two independent experiments shown.

670 F: Volcano plot shows that the majority of differentially methylated loci in *in vitro* generated
671 CAFs are hypomethylated

672 G: The differentially methylated regions in *de novo* generated CAFs are distributed throughout
673 the genome as shown in the circos plot.

674 H: Differentially methylated regions between MSC and *de novo* generated CAFs are present
675 throughout the genome and mirror the distribution of HpaII sites in the genome. Hypo refers to
676 Hypomethylated DMRs in CAFs. Hyper refers to Hypermethylated DMRs in CAFs when
677 compared to MSCs. Genome refers to distribution of HpaII loci in the genome.

678

679 **Figure 2 . CXCR4 is demethylated and overexpressed in CAFs and increases pancreatic**
680 **cancer invasiveness:**

681 A: A set of 120 transcripts are commonly hypomethylated in primary patient-derived CAFs and
682 *de novo* generated CAFs and includes *CXCR4*.

683 B,C: The *CXCR4* promoter is demethylated in primary patient-derived CAFs as seen by the
684 HELP assay (B) and quantitative MassArray Epityper analysis (C).

685 D - F: *CXCR4* knockdown in *de novo* CAFs leads to abrogation of the increased invasion of
686 Panc1 cells on co-culture. (N=3, P Value<0.05)

687 G: Co-culture with *de novo* CAFs leads to increased transwell invasion by Panc-1 cells, that is
688 abrogated after treatment of CAFs with *CXCR4* inhibitor AMD-3100 (N=3, P Value<0.05)

H: Gene expression profiling of CAFs with CXCR4 knockdown reveals significantly downregulated (*left panel*) and upregulated (*right panel*) transcripts. Salient examples of downregulated transcripts include interleukins IL-8 and IL-33 and the chemokine CCL2.

Figure 2- Figure Supplement 1: Inhibition of CXCR4 reduces PDAC cell invasion in *de novo* generated CAFs :

A: CXCR4 siRNAs lead to knockdown of protein in CAFs when compared to control siRNAs. Western blot and image intensity is shown.

B: *CXCR4* knockdown in dn-CAFs leads to abrogation of the increased invasion of Pa03C PDAC cells observed on co-culture. (N=3, *P* Value<0.05)

C: *CXCR4* knockdown using a second set of siRNAs in dn-CAFs leads to abrogation of the increased invasion of Panc1 PDAC cells observed on co-culture. (N=3, *P* Value<0.05)

D: Co-culture with dn-CAFs leads to increased transwell invasion by Pa03C PDAC cells, which is abrogated after treatment of dn-CAFs with CXCR4 inhibitor AMD-3100 (N=3, *P* Value<0.05)

Figure 3 . Increased 5hmC is seen in *de novo* CAFs generated from MSCs and is regulated by a paracrine lactate mediated metabolic flux:

A: Mapping of carbon atoms transition using U-¹³C₃-lactate. Mass isotopomer distribution (MID) shows uptake of labelled lactate by MSCs cultured in α -MEM leading to the generation of α -KG (M2 α KG) via the Krebs cycle.

B: Conditioned media (CM) from mock treated and LDH inhibitor (FX11) treated Panc-1 (PDAC) cells was collected. MSCs were exposed to media alone, control conditioned media (PDAC CM), or LDH inhibitor treated conditioned media for 14 days in order to generate *de novo* CAFs (dn-CAF). TET enzymatic activity increases in dn-CAFs after exposure to PDAC CM and is abrogated after exposure to CM from LDH inhibitor treated Panc-1 cells (N=2, *P*<0.05).

C: Quantitative analysis of 5-hmC levels by LC-MS demonstrates significant increase within dn-CAFs after treatment with Panc-1 CM, that is abrogated after exposure to CM from LDH inhibitor treated Panc-1 cells (N=2, P<0.05).

D: CAFs were exposed to exogenous lactate and α -KG levels were analysed by ELISA. (N=2, P<0.05)

E: 5hmC analysed in the resulting dn-CAFs by LC-MS. After a 2 week exposure to exogenous lactate, significantly increased 5hmC is observed in the dn-CAFs (N=2, P<0.05).

F: Schematic model of lactate flux from tumor cells to MSCs during CAF differentiation, leading to α KG generation, TET activation and conversion of 5mC into 5hmC

Figure 3- Figure Supplement 1: Metabolite tracking in CAF cells

A: Mass Isotopologue Distribution of Alanine from U-13C-Lactate tracing experiments indicate that lactate-derived carbon contributes significantly to de novo alanine synthesis. These experiments were performed in RPMI 1640 media, which does not contain alanine to highlight the de novo alanine synthesis.

B: Mole percent enrichment ratio of α -ketoglutarate to citrate indicates a net carbon contribution of lactate to α -ketoglutarate. This effect is significantly increased in culture conditions where glucose and glutamine are limited.

C, D: Ratio of intracellular levels of serine to pyruvate (left) and glycine to pyruvate (right) represent the contribution of lactate derived carbon to the serine-glycine one-carbon metabolism and subsequently to SAM cycle. In the presence of lactate, this glycolytic flux to SAM cycle is decreased. *, p<0.05; **, p<0.01 Statistical significance tested using unpaired one-tailed t-test

E: Ratio of intracellular levels of 2-hydroxyglutarate to α -ketoglutarate were previously found to be an indicator of TET activity. Lower ratios in presence of lactate in the culture media indicate an increase in TET activity. *, p<0.05; **, p<0.01 Statistical significance tested using unpaired one-tailed t-test

F: Higher ratio of intracellular levels of fumarate to α -ketoglutarate was previously found to inhibit TET activity. Lower ratios in presence of lactate in the culture media indicate an decrease in TET inhibition. *, $p < 0.05$; **, $p < 0.01$ Statistical significance tested using unpaired one-tailed t-test

Figure 3- Figure Supplement 2: Lactate increases TET activity in CAFs:

A: Conditioned media from ctrl and LDH siRNA mediated knockdown PANC-1 (PDAC) cells was collected. MSCs were exposed to control conditioned media (ctrl), or LDH knockdown conditioned and used to measure expression of CAF related genes (Fibroblasts specific protein, FSP1 (*S100A4*); α smooth muscle actin (α SMA (*ACTA1*)); and vimentin (VIM) by qRTPCR. (N=3, $P < 0.05$).

B: TET activity was analysed in MSCs exposed to exogenous lactate in dose dependant fashion (1mM to 10mM). (N=3, $P < 0.05$).

C: : TET activity was analysed in CAFs exposed to exogenous lactate (10mM) with and without Pyruvate carrier inhibitor UK5099 (5uM). (N=3, $P < 0.05$).

D: 5hmC and 5mC was analysed by ELISA in CAFs exposed to exogenous lactate with and without Pyruvate carrier inhibitor UK5099 (5uM). (N=2, $P < 0.05$).

E: TET activity was analysed in CAFs exposed to exogenous cell permeable α KG (2mM). (N=2, $P < 0.05$).

Figure 4 : 5hmC gains are seen during MSC to CAF conversion and occur at gene associated with cancer associated fibroblasts:

A: Genome wide 5hmC analysis was done by OXBS in MSCs and converted CAFs. 5hmC gains are depicted as percentage gained (CAF/MSC) and were seen to occur throughout the genome

B: 5hmC gains in CAFs were seen to occur at important genetic pathways

C-G: 5hmC gains occurred with corresponding decrease in 5mC at genes related to CAF conversion (FSP1 (*S100A4*), α SMA (*ACTA1*)); Collagen production (COL3A1), CXCR4, and the SMAD2.

Figure 5: Increased 5hmC and CXCR4 expression is seen in primary human PDAC and murine KPC PDAC tumors

A, B: 5hmC (A) and CXCR4 (B) immunohistochemical staining was done on human PDAC TMAs and grading of intensity of stain in the tumor stromal CAFs was estimated.

C: Most CAF like cells in PDAC samples were positive for 5hmC and CXCR4 (1+ to 3+ intensity). Total PDAC samples examined for 5hmC were 254 and for CXCR4 were 261.

D,E: 5hmC (D) and CXCR4 (E) immunohistochemical staining was done on mouse PDAC tumors obtained from 3 KPC mice. CAF like stromal cells were positive for 5hmC and CXCR4.

F: Single cell RNA-seq from low grade intraductal papillary mucinous neoplasm (LG IPMN), high grade intraductal papillary mucinous neoplasm (HG IPMN) and frank pancreatic ductal adenocarcinoma (PDAC) was conducted and shown in tSne plot.

G,H: Stromal cell populations cluster distinctly and are positive for α -SMA (*ACTA1*) and fibroblast activated protein.

I: Most of the stromal cells are seen in High grade IPMN and PDAC samples

J,K: CXCR4 expression is seen in 14/181 (8%) stromal cells and correlates with cells with higher collagen expression, seen in activated CAF phenotypes (T-Test, P Vccal=0.02).

Figure 6 . TET2 inactivation in MSCs leads to reduced tumor growth and inefficient CAF conversion *in vivo*:

A: Wildtype and TET2 KO MSCs were exposed to conditioned media from KPC PDAC cells for 14 days. TET2 KO MSCs demonstrated decreased fibroblastic appearance when compared to wildtype controls.

B,C: MSCs from TET2 KO treated with conditioned media from KPC PDAC cells led to significantly less KPC invasion in matrigel assays when compared with wildtype controls (N=3 independent experiments, *P* value<0.05)

D: MSCs from TET2 KO mice and WT controls were co-injected subcutaneously with KPC PDAC cells in NSG mice. Tumor growth was significantly reduced in mice injected with TET2 KO stromal cells (N=9 (TET2 KO/KPC) and N=7 (WT/KPC), *P* Value <0.05)

E,F: Tumor sizes (B) and weights (C) are significantly reduced from KPC allografts co-injected with TET2KO MSCs (*P* Value <0.05)





















G,H: Number of αSMA (*ACTA1*) positive cancer associated fibroblasts are reduced in KPC allografts co-injected with TET2KO stromal cells when compared to KPC allografts co-injected with control MSCs (N=4 in each group, *P* Value<0.01)

Supplementary File 1 : Pathways that are overexpressed and demethylated in pancreatic cancer associated fibroblasts

Supplementary File 2: Common sets of genes epigenetically altered in primary and de novo generated CAFs

Supplementary File 3: Differentially expressed transcripts after CXCR knockdown in CAF cells

813 **Table 1:** Transcription factor sites that are enriched around 5hmC gains during CAF conversion.

Rank	Name	P-value	q-value (Benjamini)	# Target Sequences with Motif	% of Targets Sequences with Motif	# Background Sequences with Motif	% of Background Sequences with Motif	Motif
1	Atf3(bZIP)/GBM-ATF3-ChIP-Seq(GSE33912)/Homer	1.00E-07	0	344	7.86%	2593.1	5.81%	
2	1(bZIP)/ThioMacPU.1-ChIP-Seq(GSE21512)/Homer	1.00E-06	0	383	8.75%	2995.7	6.71%	
3	AP1(bZIP)/K562-cJun-ChIP-Seq(GSE31477)/Homer	1.00E-05	0.0003	112	2.56%	722.1	1.62%	
4	BMAL1(bHLH)/Liver-Bmal1-ChIP-Seq(GSE39860)/Homer	1.00E-05	0.0003	857	19.57%	7567	16.95%	
5	FosI2(bZIP)/3T3L1-FosI2-ChIP-Seq(GSE56872)/Homer	1.00E-05	0.0003	149	3.40%	1040.8	2.33%	
6	BATF(bZIP)/Th17-BATF-ChIP-Seq(GSE39756)/Homer	1.00E-05	0.0003	322	7.35%	2570.8	5.76%	
7	Fra1(bZIP)/BT549-Fra1-ChIP-Seq(GSE46166)/Homer	1.00E-04	0.0006	266	6.07%	2089.9	4.68%	
8	Smad2(MAD)/ES-SMAD2-ChIP-Seq(GSE29422)/Homer	1.00E-04	0.0011	652	14.89%	5725.2	12.83%	
9	NPAS2(bHLH)/Liver-NPAS2-ChIP-Seq(GSE39860)/Homer	1.00E-04	0.0026	543	12.40%	4734.1	10.61%	
10	Smad4(MAD)/ESC-SMAD4-ChIP-Seq(GSE29422)/Homer	1.00E-03	0.003	655	14.96%	5819.3	13.04%	
11	Smad3(MAD)/NPC-Smad3-ChIP-Seq(GSE36673)/Homer	1.00E-03	0.0035	1238	28.27%	11533	25.84%	
12	Usf2(bHLH)/C2C12-Usf2-ChIP-Seq(GSE36030)/Homer	1.00E-03	0.0042	200	4.57%	1574.5	3.53%	
13	HIF2a(bHLH)/785_OHIF2a-ChIP-Seq(GSE34871)/Homer	1.00E-03	0.0066	178	4.06%	1393.5	3.12%	
14	CEBP:CEBP(bZIP)/MEF-Chop-ChIP-Seq(GSE35681)/Homer	1.00E-03	0.0141	107	2.44%	789.3	1.77%	
15	MafA(bZIP)/Islet-MafA-ChIP-Seq(GSE30298)/Homer	1.00E-02	0.023	315	7.19%	2709.1	6.07%	
16	Bach2(bZIP)/OCILy7-Bach2-ChIP-Seq(GSE44420)/Homer	1.00E-02	0.023	86	1.96%	622.7	1.40%	
17	Brachyury(T-box)/Mesoendoderm-Brachyury-ChIP-Seq(GSE54963)/Homer	1.00E-02	0.0326	134	3.06%	1056.7	2.37%	
18	USF1(bHLH)/GM12878-Usf1-ChIP-Seq(GSE32465)/Homer	1.00E-02	0.0387	190	4.34%	1574.4	3.53%	
19	n-Myc(bHLH)/mES-nMyc-ChIP-Seq(GSE11431)/Homer	1.00E-02	0.0511	272	6.21%	2354.1	5.27%	
20	Flt1(ETS)/CD8-Flt1-ChIP-Seq(GSE20898)/Homer	1.00E-02	0.0528	439	10.03%	3953.7	8.86%	

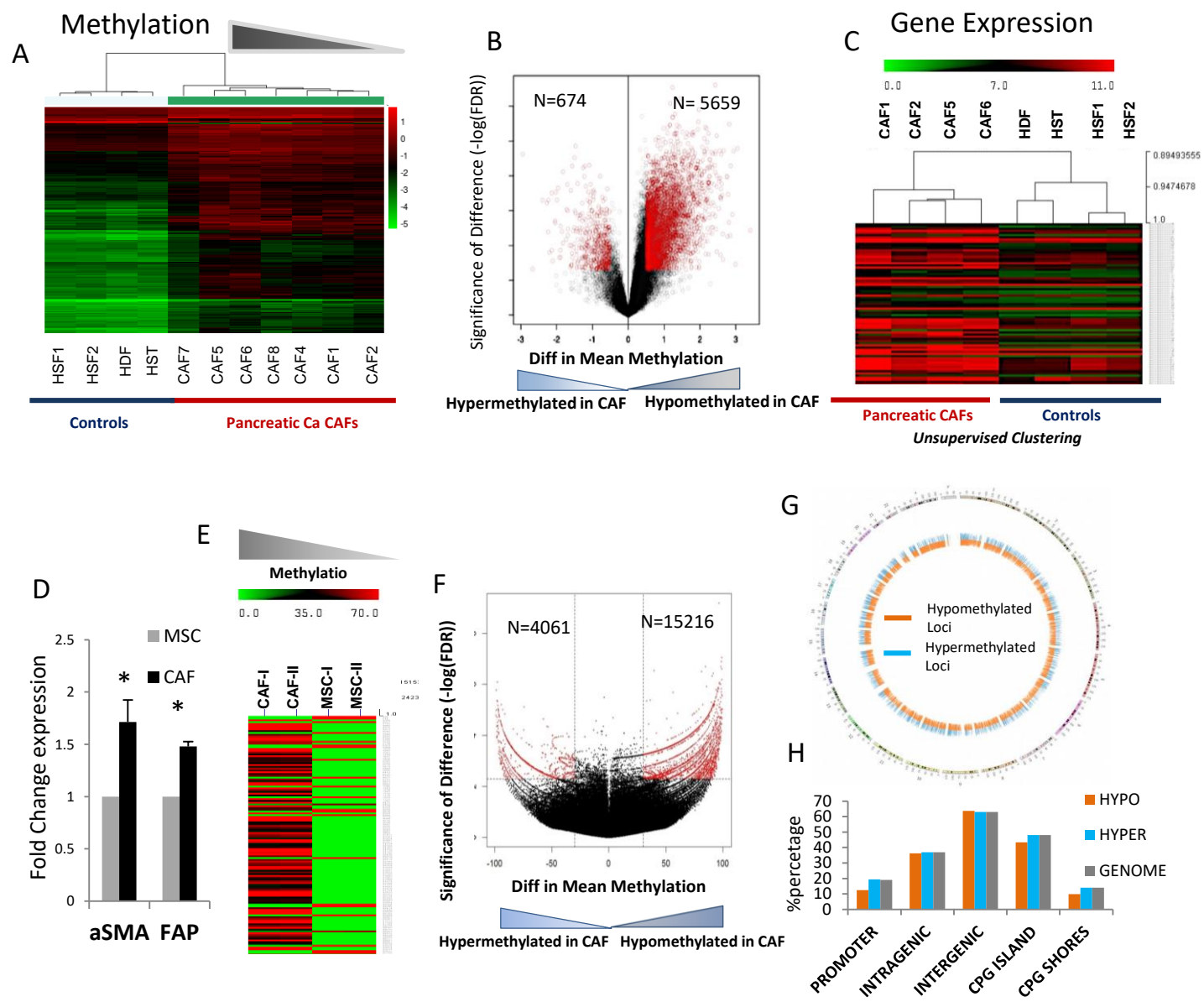


Fig 1

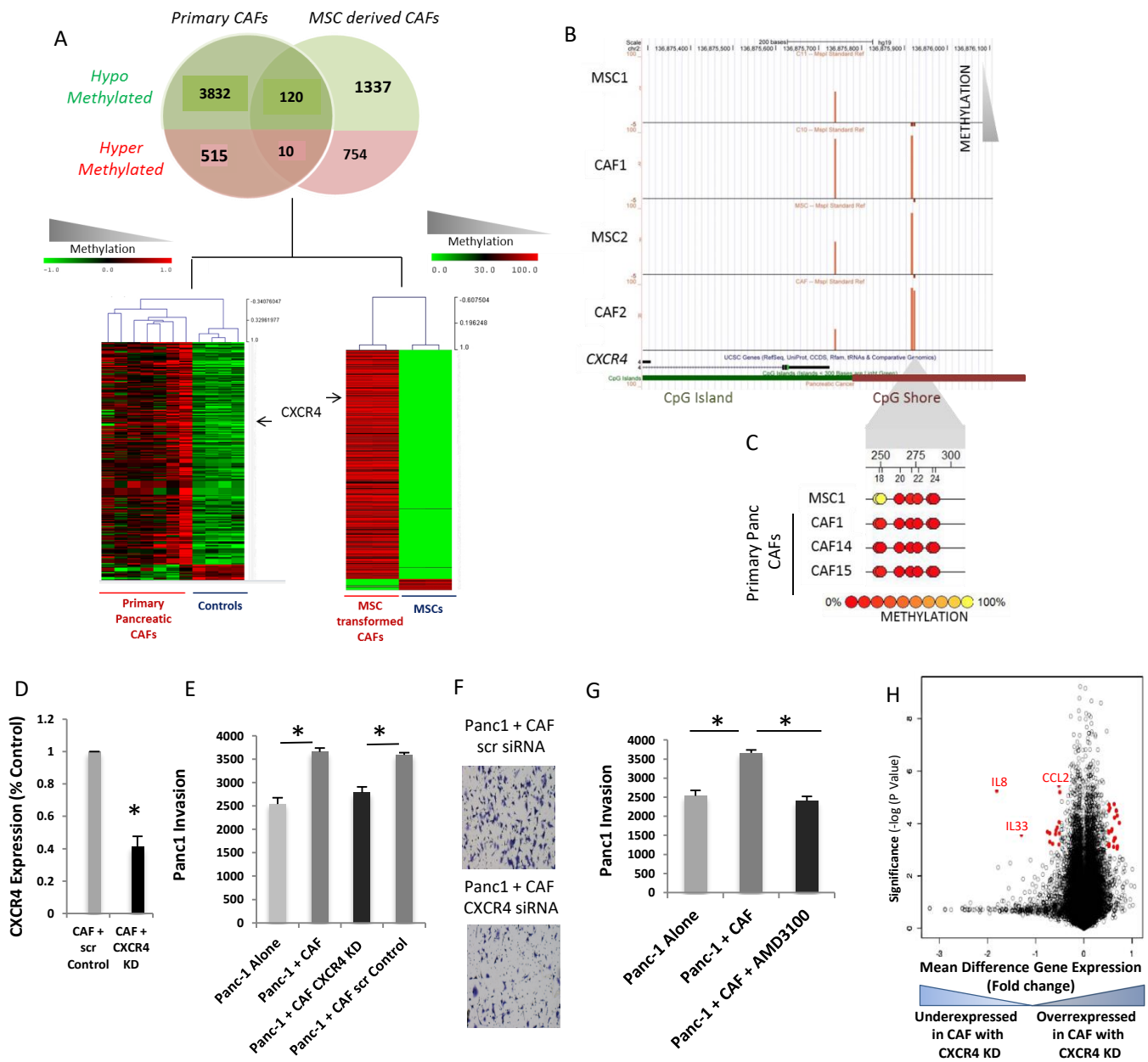


Fig 2

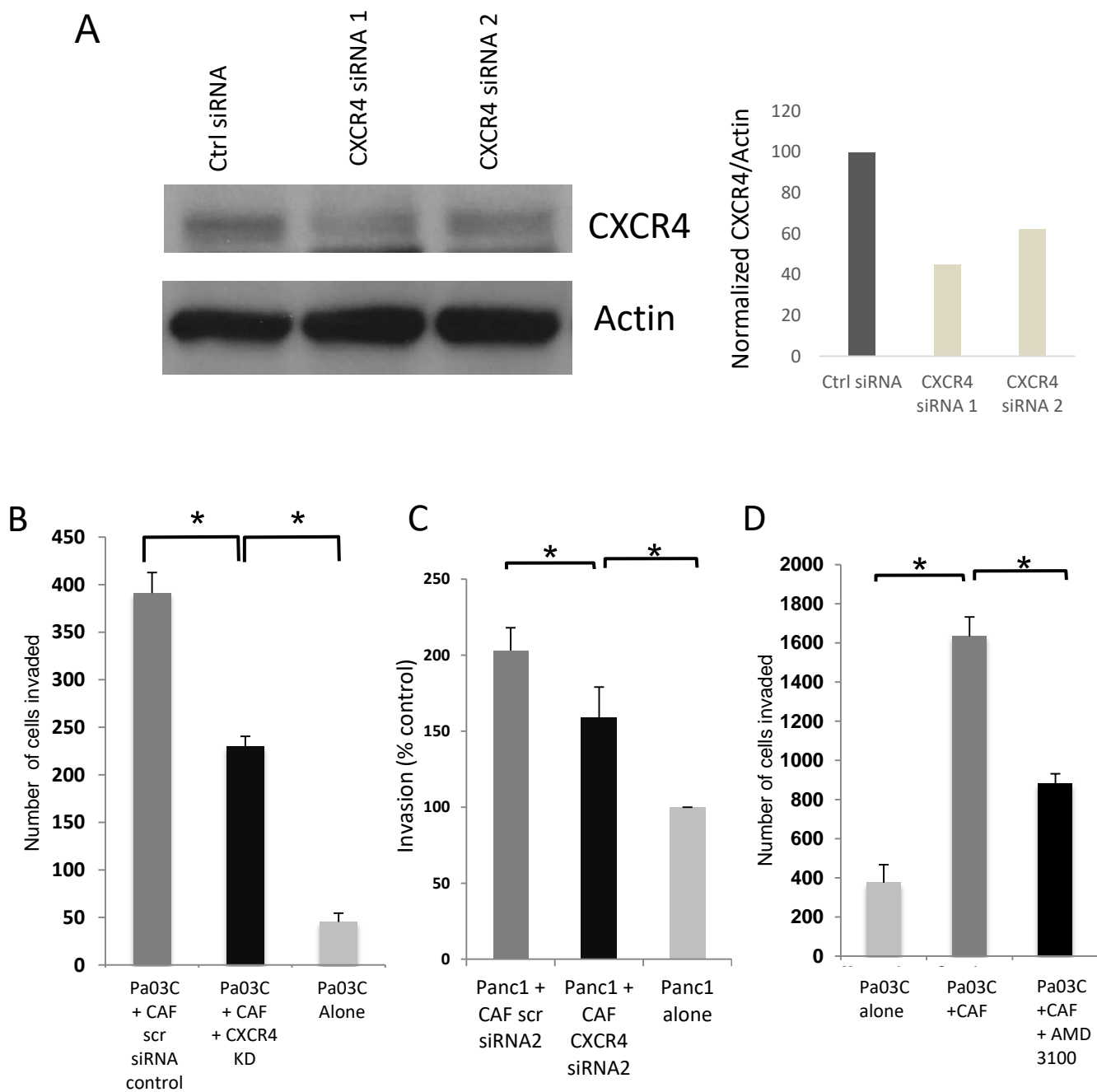


Figure 2- Figure Supplement 1

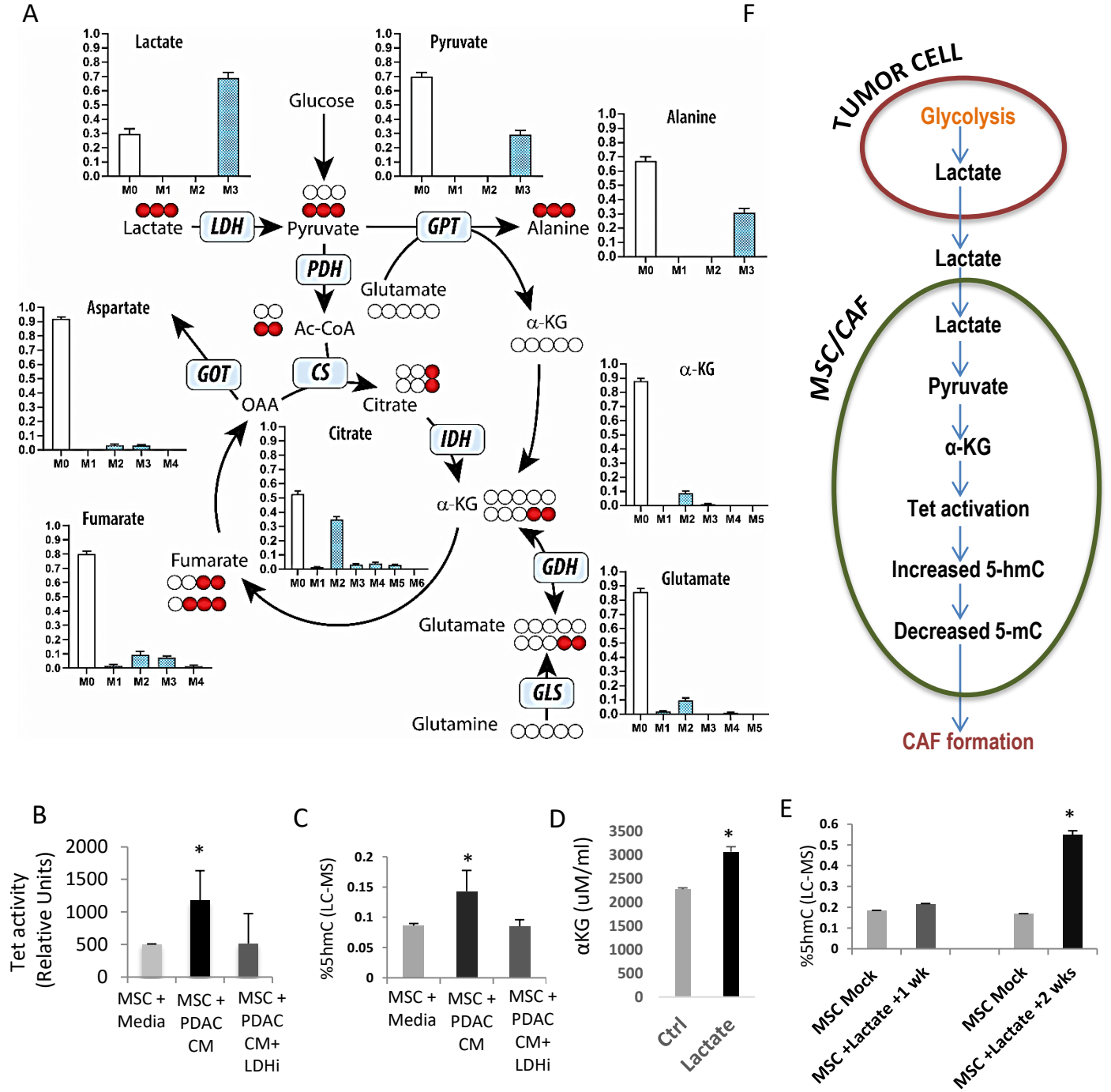


Fig 3

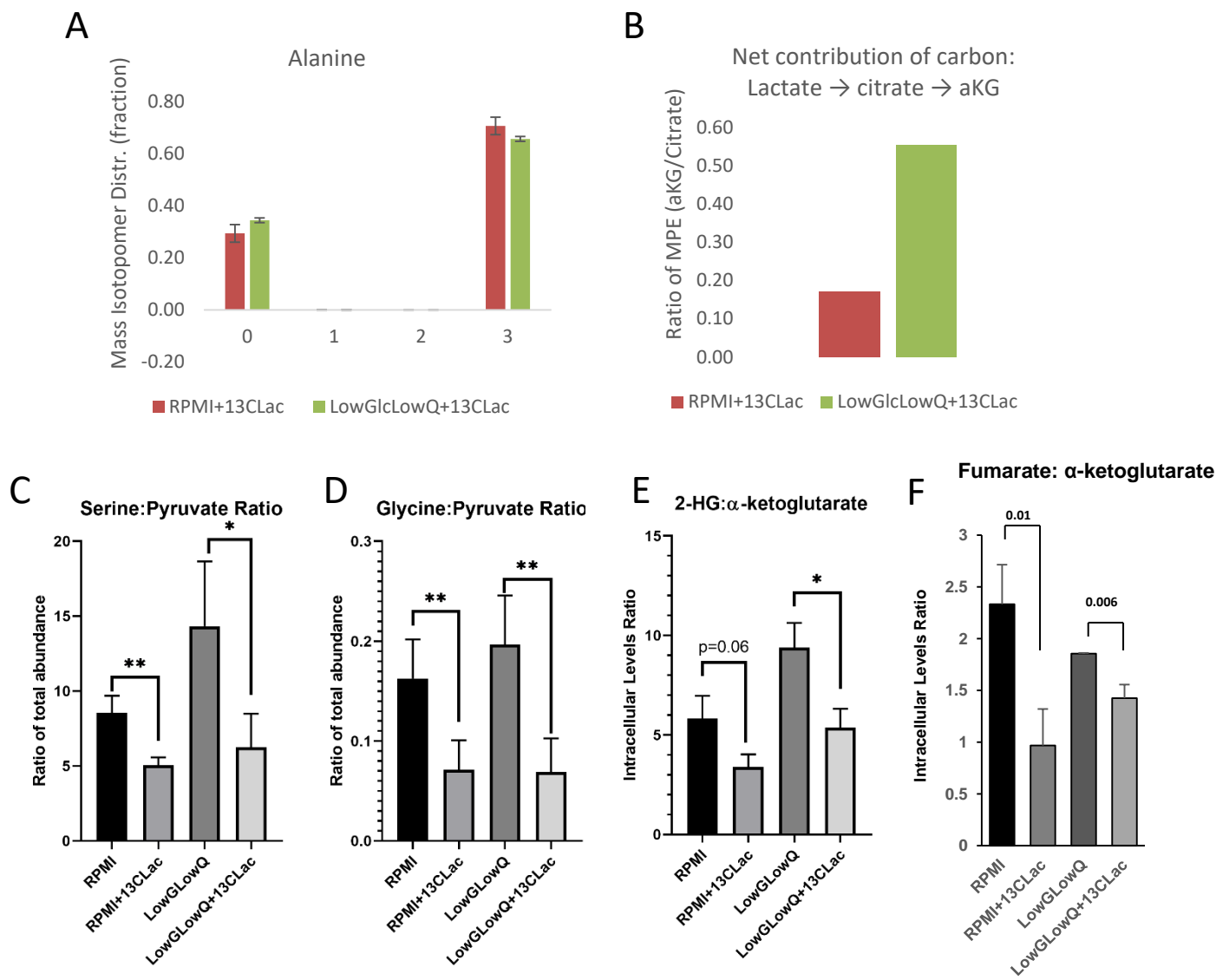


Figure 3- Figure Supplement 1

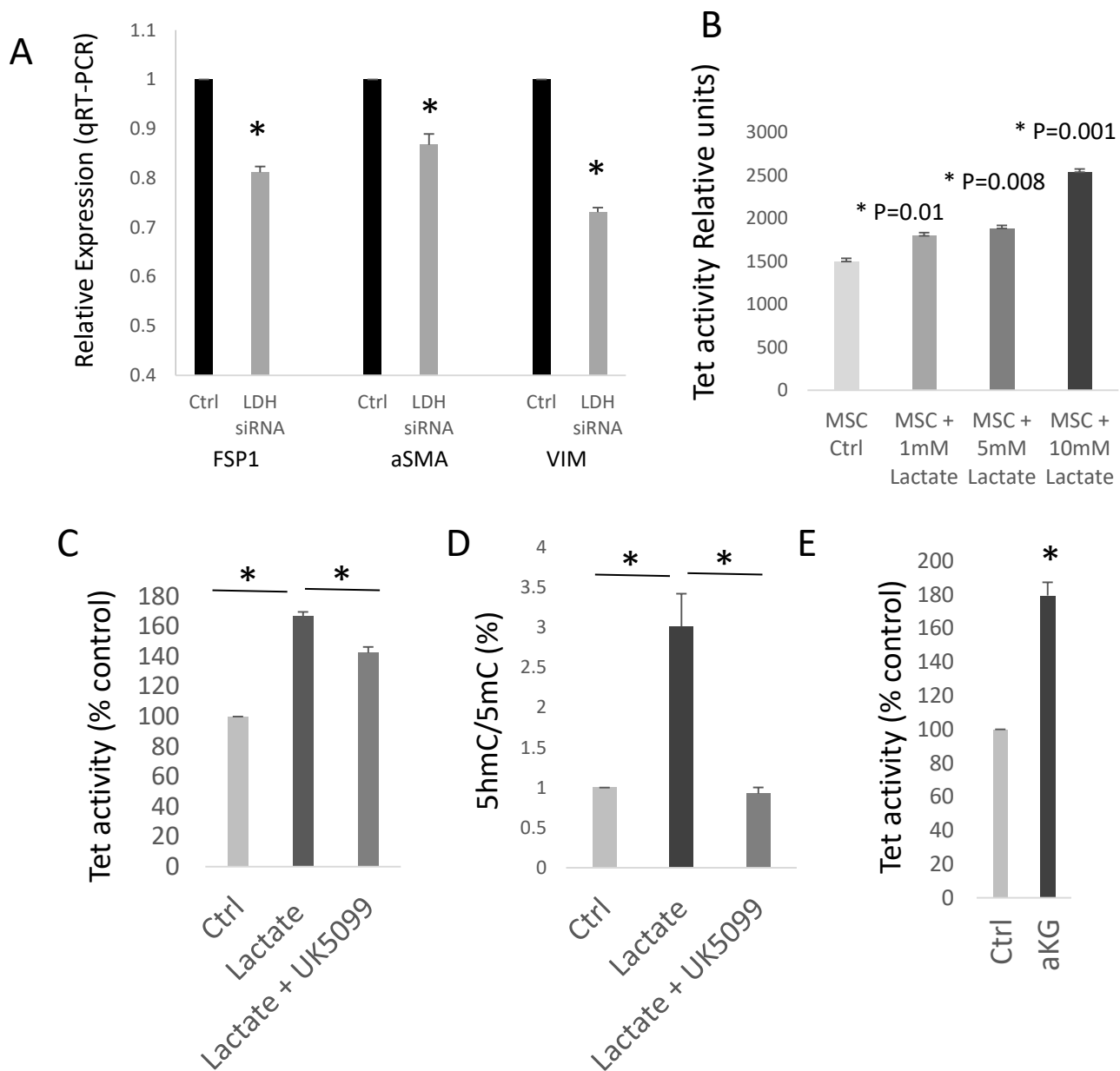


Figure 3- Figure Supplement 2

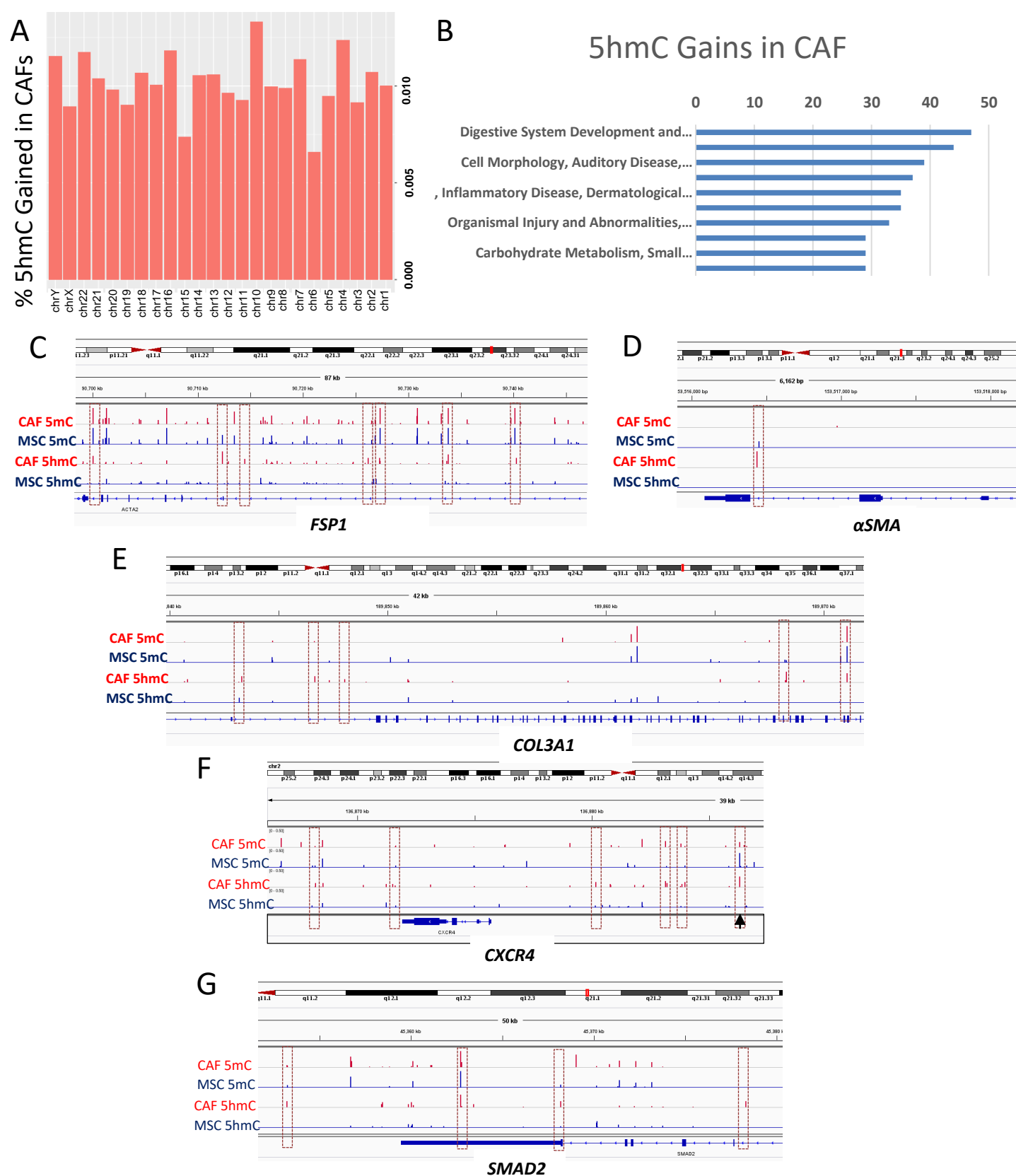


Fig 4

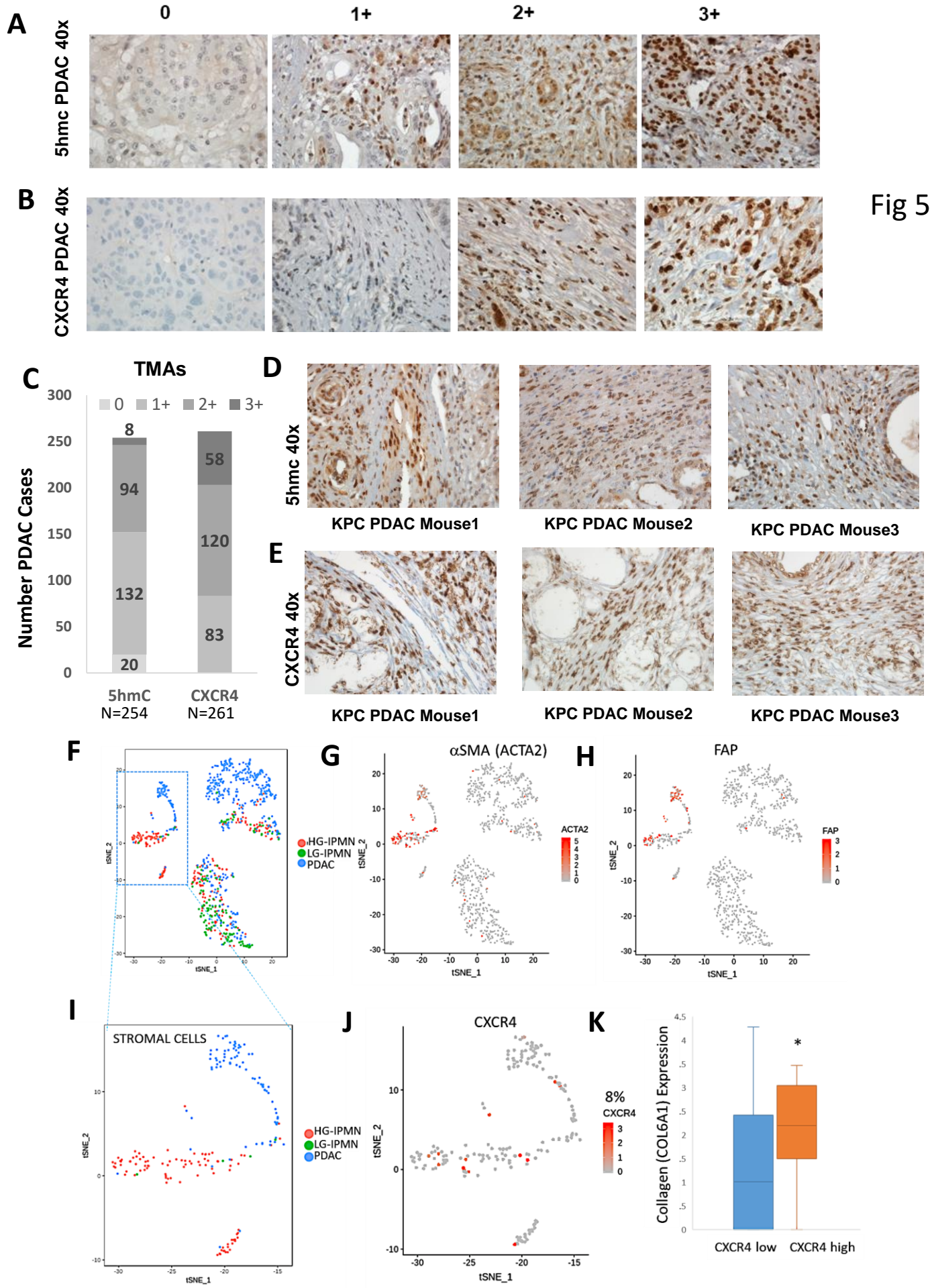


Fig 5

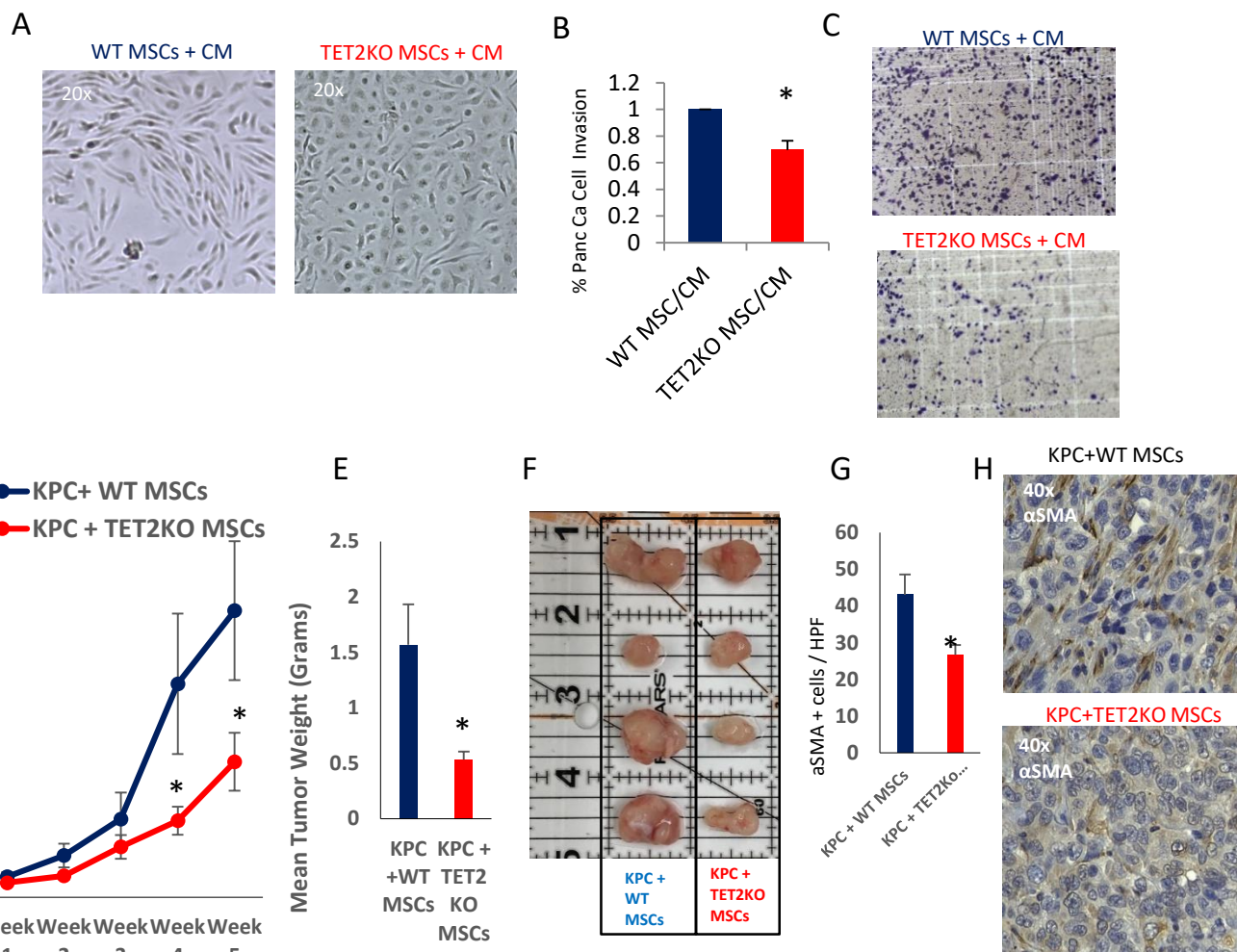


Fig 6