Supplemental Methods

In silico analysis of PAF1C subunits, PHF5A, and DDX3 in PC data sets

The GEPIA web server (http://gepia.cancer-pku.cn/) was used to analyze TCGA PC expression data. Herein, expression data for PAF1, CTR9, LEO1, CDC73, SKI8, PHF5A, and DDX3 were obtained. The data were downloaded using the web portal and subsequently plotted[1]. The normalized PAAD gene expression data (FPKM-UQ) was obtained from Genomic Data Commons (GDC) web portal (portal.gdc.cancer.gov). Grouped boxplots of PAF1, other PAF1C subunits, DDX3, and PHF5A were made using the ggplot2 package in R. The boxplots were grouped based on the ABSOLUTE purity classifications, as provided by Raphael et al., 2017[2].

Orthotopic tumor implantation

SW1990, Capan1, and CD18 cell lines were engineered for inducible PAF1 knockdown using a lentiviral system; these cell lines were injected in the pancreas of 5-week-old athymic nude mice orthotopically as described previously[3]. Doxycycline hyclate (Sigma-Aldrich) was administered to mice assigned to the PAF1 KD group via drinking water at a concentration of 2 mg/ml in 2% sucrose solution. Control mice received 2% sucrose in drinking water.

Subcutaneous tumor implantation

Control and PAF1 KD cells (40,000) were suspended in 500 µl of PBS and mixed with Matrigel (Fisher Scientific) in a 1:1 ratio. Cells (4,000) in 100 µl volume of PBS-matrigel mix were injected subcutaneously into the right and left flanks of 5-week-old athymic nude mice. Four mice were used per group, and the appearance of tumors was checked by palpitation two times per week. Animals were sacrificed soon after the tumor nodule reached 0.8 cm diameter.

Determination of tumor recurrence *in vivo*

Tumor recurrence has previously been evaluated in mouse models of PC[4]. Briefly, 0.5×10^6 luciferase labelled SW1990 cells were orthotopically implanted in 6 week-old nude mice. Tumor formation was ascertained using IVIS. Two weeks post-implantation of cancer cells in the pancreas, tumors were resected as described previously[4]. The primary tumor in the pancreatic

tail was carefully removed by disconnecting the pancreatic body, including blood vessels, by using one clip followed by coagulation and cutting. We confirmed negative margins after the resection to ensure complete tumor removal. Six days after tumor resection, mice were imaged using IVIS to assess tumor burden, and mice were then randomized into two groups: control and PAF1 KD. Mice segregated into PAF1 KD received doxycycline (2 mg/ml) in 2% sucrose in drinking water, and the other group received 2% sucrose only (vehicle control). Tumor-bearing mice in both groups were followed for three weeks and tumor recurrence was evaluated using IVIS and histological analysis at the endpoint.

Limiting dilution assay

Cell suspension of Capan1 cells was prepared for five different concentrations (C1 = 100,000 cells; C2 = 50,000 cells; C3 = 10,000 cells; C4 = 1000 cells; C5 = 100 cells) in a 50 µl volume of PBS. Cells were mixed with 50 µl Matrigel and were subcutaneously implanted into the right and left flanks of 5-week-old athymic nude mice. Mice were segregated into two groups: control and PAF1 KD. Three mice were used per group per cell concentration (6 tumors in each group/cell concentration), and the appearance of tumors was checked by palpitation thrice per week. Tumor dimensions were measured using Vernier Calipers for the determination of tumor volume. Tumors were harvested when control tumors became necrotic.

Organoid generation and culture

In brief, tumor organoids were established after tumor resection, mechanical and enzymatic digestion of pancreatic tumor from KPC autochthonous mouse model with 0.012% (w/v) collagenase XI (Sigma) and 0.012% (w/v) dispase (GIBCO) in DMEM media containing 1% FBS (GIBCO) and embedded in growth factor reduced (GFR) Matrigel (BD)[5]. These organoids were maintained and cultured in complete AdDMEM/F12 medium supplemented with HEPES (Invitrogen), Glutamax (Invitrogen), penicillin/streptomycin (Invitrogen), B27 (Invitrogen), Primocin (1 mg/ml, InvivoGen), N-acetyl-L-cysteine (1 mM, Sigma), mouse recombinant Wnt3a (100 ng/ml, Milipore), human recombinant RSpondin1 (1 µg/ml, Nuvelo), Noggin (0.1 mg/ml,

Peprotech), epidermal growth factor (EGF, 50 ng/ml, Peprotech), Gastrin (10 nM, Sigma), fibroblast growth factor 10 (FGF10, 100 ng/ml, Preprotech), Nicotinamide (10 mM, Sigma), and A83- 01 (0.5 mM, Tocris). Human organoids were derived from the pancreas tissue obtained from fresh tissue bank at UNMC and cultured using the same protocol as followed with mouse organoids with the addition of Wnt3A (1X (50% V/V), Preprotech).

ALDEFLUOR Assay

The cells expressing high levels of activity of the enzyme aldehyde dehydrogenase (ALDH) were stained using ALDEFLUOR reagent (Stem Cell Technologies, Cambridge, MA, USA. Cat: 01700) using the manufacturer's instructions**.** A specific inhibitor of ALDH, diethylaminobenzaldehyde (DEAB), was used to control for background fluorescence.

Analysis of $CD44^+$ CSCs using flow cytometry

Cells were trypsinized, counted, and resuspended to get one million cells/100 µl per tube. Cells were incubated with human CD44-APC antibody (Miltenyi Biotec Inc. Auburn, CA, USA .Cat:130-098-110) for 15 minutes at 4 °C, followed by two PBS washes and resuspension in 500 µl media. Appropriate single and double color controls were included, and cells with a high expression of CD44+ (high CD44+) were gated using appropriate controls.

Apoptosis Assay

SW1990 SP and 9-26 NP cells were treated with vehicle control DMSO or RK-33 (5M) for 48 h. Cell supernatant and adherent cells were collected, stained with annexin V and propidium iodide (PI), and analyzed by flow cytometry as described previously[6].

Immunoblot Assay

Cells were processed for protein isolation and Western Blotting using standard procedures, as described previously [7]. The following primary antibodies were used: anti-PAF1, anti-LEO1, anti-CDC73, anti-CTR9 (Bethyl Laboratories, Montgomery, TX, USA); anti-PHF5A (Proteintech, Rosemont, Illinois, USA); anti-OCT3/4, anti-SOX-2, anti-CD24, anti-ESA (Santa Cruz Biotechnology, Dallas, TX, USA); anti-CD44 (Cell signalling Technology, Danvers, MA, USA);

anti-CD133 (Abnova, Walnut, CA, USA); anti–β-Catenin (Sigma); anti-DDX3, anti-Lgr5, and anti-SOX-9 (Abcam, Cambridge, MA, USA). All antibodies were incubated overnight at 4 °C. β-ACTIN was used as a loading control.

Immunofluorescence analysis

Cells were plated, fixed, and processed as described previously[7]. Primary antibodies specific for mouse-PAF1 (1:100 in PBS), mouse-OCT3/4 (1:100), rabbit-β-CATENIN (1:100), mouse-CD44 (1:250), and rabbit-ESA (1:1500) were used with a 4 h incubation for cells. For human and mouse tissues, we followed the same procedure mentioned previously[8], but incubation with primary antibodies was performed overnight at 4 °C. Following primary antibody incubation, the cells and tissue sections were processed using standard procedures, as described previously[7]. Confocal images were collected using Zeiss LSM800 confocal microscope with a 63×/1.4 NA oil objective.

Quantification of fluorescent intensity: Multichannel snapshots were split into separate channels in ImageJ. A region of interest was drawn around individual cells in one of the two channels using the “freehand” tool. This region was then subjected to the measure plugin, and intensity was calculated by multiplying area and mean intensity. The measurement of intensity was made for 10 cells in a field and 5-6 images per condition. The intensity values were represented relative to the control condition in arbitrary units (A.U.).

Colony formation assay (clonogenic assay)

SW1990 SP PD2/PAF1 CRISPR KD cells and control SW1990 SP cells were trypsinized and seeded at a density of 1000 cells/well in a 6-well plate in triplicates. The cells were cultured in CSC-specific media, with media changed once in two days. After two weeks of growth, the cells were fixed with 100% methanol and stained with crystal violet stain (0.1%, w/v in 20 nm 4-morpholinepropanesulfonic acid; Sigma) before the colonies started to merge. Total colony area was quantified using ImageJ.

Tumor sphere assay

Capan1 SP cells that were engineered to knockdown PAF1 using Dox-induced shRNA were seeded in triplicates in a 96-well non-adherent plate (Corning Inc., Corning, New York, USA) in CSC-specific media at a concentration of 100 cells/well in two sets. The cells in suspension culture were observed under the microscope and fresh media was added every alternate day without removing the existing media in the set designated as control, whereas the set designated as KD set received Dox at 2 μg/ml every alternate day. Treatment with RK-33 was initiated 24 h after cell seeding. A week later, multiple images were taken per well for different fields of view. The diameter of each tumor sphere was measured using Motic Images Plus 2.0 ML software; the plot depicting the diameter of tumor sphere in PAF1 KD cells and control cells was plotted using MedCalc software.

Cell motility assay

The motility assay was performed by using a chamber containing a monolayer-coated polyethylene teraphthalate membrane (six-well insert, a pore size of 8 μm; Becton Dickinson, Franklin Lakes, NJ, USA). Both control and PAF1 knockdown (Dox treated) SW1990 cells (1×10^6 cells per well) were seeded in six-well plates. After a 24 h incubation period, the migrated cells that had reached the lower chamber were stained with a Diff-Quick stain set and counted in different fields. The average number of migrated cells per representative field was calculated.

Wound-healing assay

The wound-healing assay was performed as described previously[9]. Briefly, 4×10^6 control and PAF1 KD (2 mg/ml doxycycline treated) cells were seeded in 6-well plates. After 24 h, a scratch was made on the bottom of each well using a P200 pipette tip. Photographs of the scratch were taken at 0 h and 24 h. The pictures were then analyzed to quantify the change in size of the scratches as the cells migrated into the space; wound closure was compared between control and PAF1 depleted cells.

Immunoprecipitation analysis

SW1990 SP, HPDE, and F9 cells were cultured in their respective culture conditions, and the lysates were collected in a CHAPS non-denaturing immunoprecipitation buffer (20 mM Tris, pH 7.5, 200 mM NaCl, 1%). Immunoprecipitation was performed with PAF1 (ChIP-grade, Abcam) and anti-Phf5a (Proteintech) antibodies using Dynabeads protein G (Invitrogen). Following pre-clearing, 500 μg of lysates prepared at 1 mg/ml were incubated with 3 μg of respective antibodies overnight in a rotor at 4 °C. The next day, dyna beads (20 μl per reaction) were added to the antibody-lysate mix and incubated for rotation at 4 °C for 6 h. Thereafter, the unbound antibodies were washed with IP wash buffer (40 mM HEPES pH 7.4, 500 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, and 1% Triton X-100), and the immunoprecipitated proteins were eluted using a high pH ammonium buffer and were concentrated using a speed Vac. The immunoprecipitates, or total cell lysates, were transferred onto the PVDF membrane after being resolved on 10% SDS PAGE, and thereafter were incubated overnight at 4 °C with primary antibodies (anti-PAF1, anti-OCT3/4, anti-LEO1, anti-CTR9, and anti-CDC73).

RNA isolation and quantitative reverse-transcription Polymerase Chain Reaction (qRTPCR)

Total cellular RNA was extracted from cells using the RNAeasy kit (Qiagen, Hilden, Germany) and processed for reverse transcription to generate cDNA as described previously[10]. cDNA products were assayed by quantitative real-time PCR using SYBR Green incorporation. The expression of all genes was normalized to that of internal control β-actin and expressed relative to the indicated reference sample (average ± SEM of triplicate reactions).

Human transcription factors PCR array

Total RNA isolated from SW1990 SP and PAF1-depleted SP cells were reverse-transcribed using RT^2 SYBR qPCR master mix (330,401, Qiagen). Aliquots of 25 μl of the mix of both samples were added in separate 96-well PCR array kits containing lyophilized gene-specific

primer set (PAHS-075Z, Qiagen). Threshold cycles were used to calculate fold change using the online web server of RT^2 profiler PCR array data analysis[11].

Chromatin Immunoprecipitation (ChIP)-Sequencing

SW1990 SP cells were fixed with 0.4% formaldehyde and 1.5 mM EGS (ethylene glycol bis(succinimidyl succinate)) (Thermo Scientific Fischer), washed, harvested, and resuspended in 500 µl SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl [pH 8.1], 1 mM PMSF, and 1 µg/ml aprotinin). Samples were sonicated and diluted in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl [pH 8.1], 167 mM NaCl, 1 mM PMSF, and 1 µg/ml aprotinin). For input control, 10% of sonicated samples were separated. Immunoprecipitation was performed with ChIP-grade PAF1 (Abcam) and Phf5a (Proteintech) antibody as previously described. Chromatin extracts were pulled down with protein A+G beads. The samples were washed extensively with wash buffers (low salt, high salt, LiCl, and Tris/EDTA buffers), eluted with SDS elution buffer, and subjected to reverse cross-linking and proteinase digestion. IgG antibody was used as a control for the ChIP assay.

ChIP-seq Analysis: The captured and purified DNA was prepared for high throughput sequencing using the New England Biolabs NEBNext Ultra II DNA Library Prep Kit for Illumina. The resulting indexed libraries were sequenced by the UNMC Sequencing Core Facility using an Illumina NextSeq 500 Genome Analyzer. Initial raw sequence files were processed based on the following steps. Adaptor sequences and low quality (Phred score < 20) ends were trimmed from sequences using Trim Galore software package (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Resulting fastq files were aligned to the human genome (GRCh38/hg38) using the sequence aligner Bowtie2 (version 2.2.3)[12]. The software package Picard routine, MarkDuplicates, (http://broadinstitute.github.io/picard/) was used to remove sequence duplications. For peak calling of ChIP enriched regions, the MACS2 peak caller software (version 2.1.1)[13] of each ChIP

to corresponding input DNA sample was used to determine binding regions based on a FDR adjusted p-value (q-value) <0.05. BigWig files were generated using the deeptools bamCoverage routine (https://deeptools.readthedocs.io/en/develop/). Alignment of significant peaks to gene-specific regions was accomplished using the bedtools routine intersect (https://bedtools.readthedocs.io/en/latest/). Raw files were submitted to GEO (accession number GSE144371).

RNA-Sequencing

SW1990 SP control and PAF1 CRISPR KD cells were lysed in RNA lysis buffer supplied with the mirVana RNA isolation kit (Thermo Scientific Fischer, n = 2). Total RNA was isolated in accordance with the manufacturer's protocol. Whole transcriptome analysis (RNA sequencing) was performed on PAF1 KD and control cells at the Sequencing Core Facility at the University of Nebraska Medical Center. RNA quality was checked with an Agilent Bioanalyzer (Agilent Technologies, Inc.) and all RINs (RNA integrity numbers) were 10. Library preparation was achieved using Illumina TruSeq RNA Sample Preparation Kit according to the manufacturer's protocol. Fifteen cycles of PCR amplification were performed for each library, followed by examination of size distribution using an Agilent Bioanalyzer (Agilent Technologies, Inc.) and a DNA 1000 chip. Each sample used one out of twelve unique indices (Illumina). All libraries displayed a band between 200–500 bp with a peak at approximately 290 bp. The Qubit 2.0 Fluorometer (Life Technologies, Inc.) was used to quantitate the libraries. Loading was performed at a concentration of 6 pM. Sequencing was done on HiSeq 2500 sequencer in rapid mode. It was a single read, 50 cycle, sequencing run, and onboard clustering and V2 chemistry was used. Raw files for RNA-seq were submitted to GEO (accession number GSE144371).

Mass Spectrometry

Sample prep summary: Protein fractions were excised from SDS-PAGE gel, destained, reduced with tris-carboxyethylphosphine, alkylated with iodoacetamide, and digested overnight with

sequencing-grade trypsin (Promega). Tryptic peptides were eluted from the gel and concentrated to 20 µl by vacuum centrifugation and analyzed using a high-resolution mass spectrometry nano-LC-MS/MS Tribrid system, Orbitrap Fusion™ Lumos™ coupled with UltiMate 3000 HPLC system (Thermo Scientific). 500 ng of peptides were run by the pre-column (Acclaim PepMap™ 100, 75µm × 2cm, nanoViper, Thermo Scientific) and the analytical column (Acclaim PepMap™ RSCL, 75 µm × 50 cm, nanoViper, Thermo Scientific). The samples were eluted using a 120-min linear gradient of ACN (5-45%) in 0.1% FA.

Results summary: All MS/MS samples were analyzed using Mascot (Matrix Sciences, London, UK, version 2.6.). Mascot was set up to search the SwissProt database (selected for Homo sapiens, 2018_06, 20361 entries), assuming the digestion enzyme trypsin. Parameters on MASCOT were set as follows: Enzyme: trypsin, Max missed cleavage: 1, Peptide charge: 1+, 2+, and 3+, Peptide tolerance: ± 0.8 Da, Fixed modifications: carbamidomethyl (C), Variable modifications: oxidation (M), MS/MS tolerance: ± 0.6 Da with error tolerant search, Instrument: ESI-TRAP. MASCOT results for different gel cuts of the same sample were combined and analyzed using Scaffold, which allows multiple search results to be condensed into a single result file.

Scaffold (version 4.8.7, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Peptide Prophet algorithm[14] with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm[15]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters**.**

Statistical analysis

Student's *t*-test was used to determine the statistical significance between control and PAF1 knockdown group in all the experiments pertaining to this study. Statistical analysis and the generation of graphs were performed using GraphPad Prism and MedCalc software. Alfa level of 0.05 was used for statistical significance unless otherwise indicated. Error bars were given on the basis of calculated standard error values.

Supplementary references

1. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017;45:W98-W102.
2. Cancer Genome Atlas Research Network. Electronic address aadhe, Cancer Genome Atlas Research N. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell 2017;32:185-203 e13.
3. Vaz AP, Deb S, Rachagani S, et al. Overexpression of PD2 leads to increased tumorigenicity and metastasis in pancreatic ductal adenocarcinoma. Oncotarget 2015.
4. Gürlevik E, Fleischmann-Mundt B, Brooks J, et al. Administration of Gemcitabine After Pancreatic Tumor Resection in Mice Induces an Antitumor Immune Response Mediated by Natural Killer Cells. Gastroenterology 2016;151:338-350.e7.
5. Boj SF, Hwang CI, Baker LA, et al. Organoid models of human and mouse ductal pancreatic cancer. Cell 2015;160:324-38.
6. Lakshmanan I, Batra SK. Protocol for Apoptosis Assay by Flow Cytometry Using Annexin V Staining Method. Bio Protoc 2013;3.
7. Vaz AP, Ponnusamy MP, Rachagani S, et al. Novel role of pancreatic differentiation 2 in facilitating self-renewal and drug resistance of pancreatic cancer stem cells. Br. J Cancer 2014;111:486-496.
8. Nimmakayala RK, Seshacharyulu P, Lakshmanan I, et al. Cigarette Smoke Induces Stem Cell Features of Pancreatic Cancer Cells via PAF1. Gastroenterology 2018;155:892-908 e6.
9. Lakshmanan I, Rachagani S, Hauke R, et al. MUC5AC interactions with integrin beta4 enhances the migration of lung cancer cells through FAK signaling. Oncogene 2016;35:4112-21.
10. Ponnusamy MP, Deb S, Dey P, et al. RNA polymerase II associated factor 1/PD2 maintains self-renewal by its interaction with Oct3/4 in mouse embryonic stem cells. Stem Cells 2009;27:3001-3011.
11. Kumar S, Das S, Rachagani S, et al. NCOA3-mediated upregulation of mucin expression via transcriptional and post-translational changes during the development of pancreatic cancer. Oncogene 2015;34:4879-89.
12. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357-9.
13. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9:R137.
14. Keller A, Nesvizhskii AI, Kolker E, et al. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 2002;74:5383-92.
15. Nesvizhskii AI, Keller A, Kolker E, et al. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 2003;75:4646-58.

Figure S1

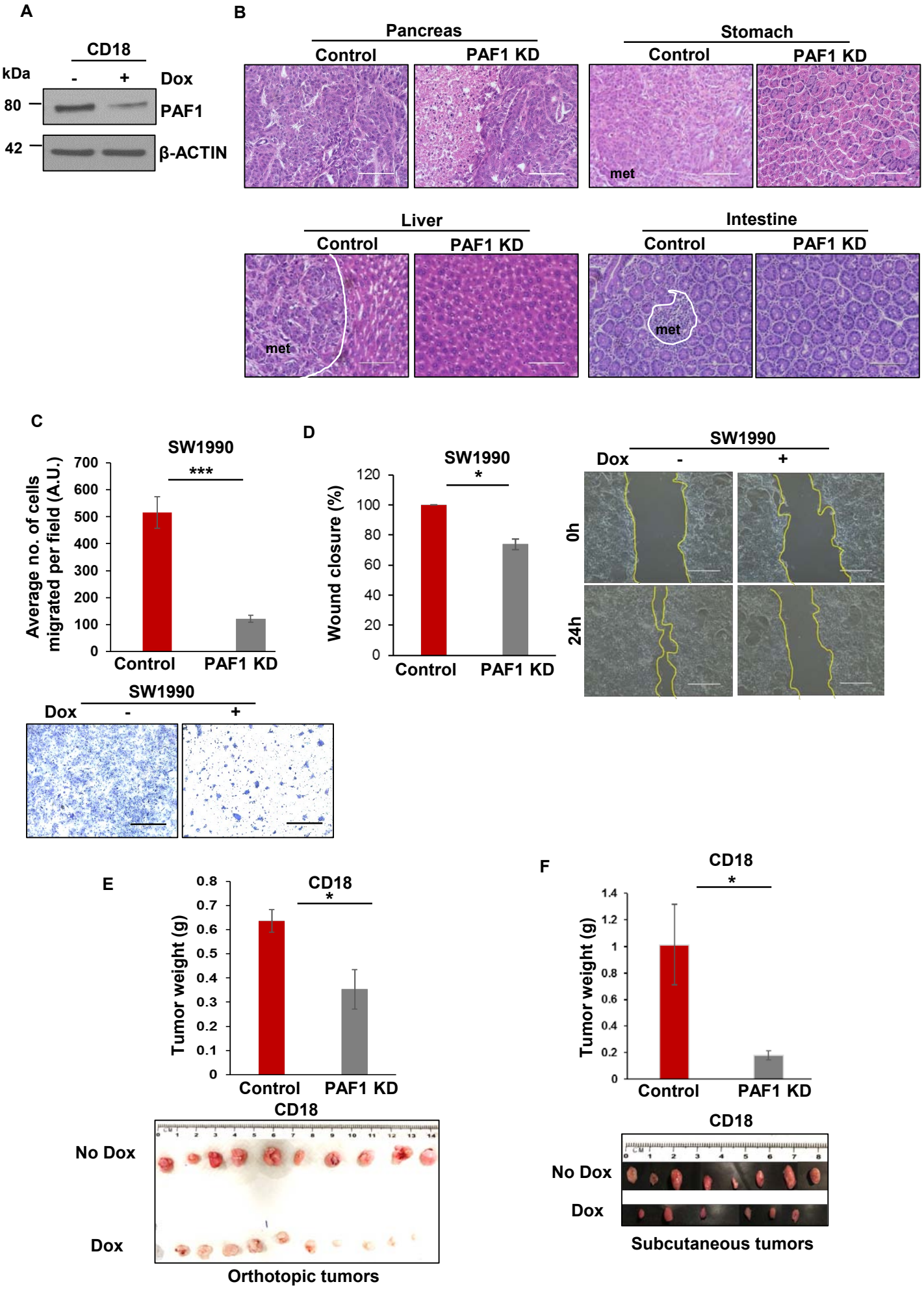

A
CD18
kDa
- +
Dox
80
PAF1
42
β-ACTIN
B
Pancreas
Control
PAF1 KD
Stomach
Control
PAF1 KD
met
Liver
Control
PAF1 KD
met
Intestine
Control
PAF1 KD
met
C
SW1990
***
Average no. of cells migrated per field (A.U.)
700
600
500
400
300
200
100
0
Control
PAF1 KD
SW1990
Dox
- +
D
SW1990
*
Wound closure (%)
120
100
80
60
40
20
0
Control
PAF1 KD
SW1990
Dox
- +
0h
24h
E
CD18
*
Tumor weight (g)
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
Control
PAF1 KD
CD18
No Dox
Dox
Orthotopic tumors
F
CD18
*
Tumor weight (g)
1.4
1.2
1
0.8
0.6
0.4
0.2
0
Control
PAF1 KD
CD18
No Dox
Dox
Subcutaneous tumors

Figure S1

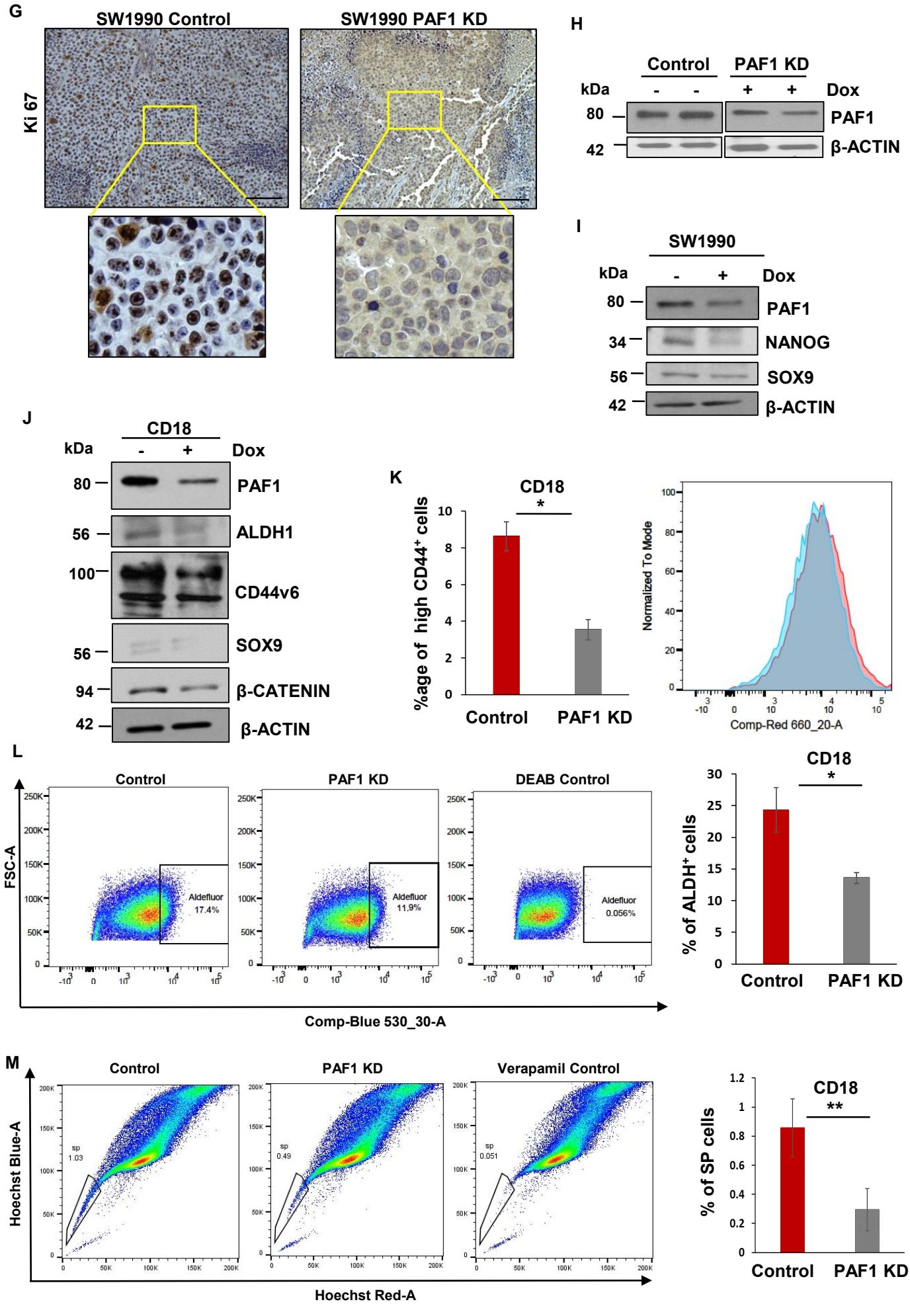
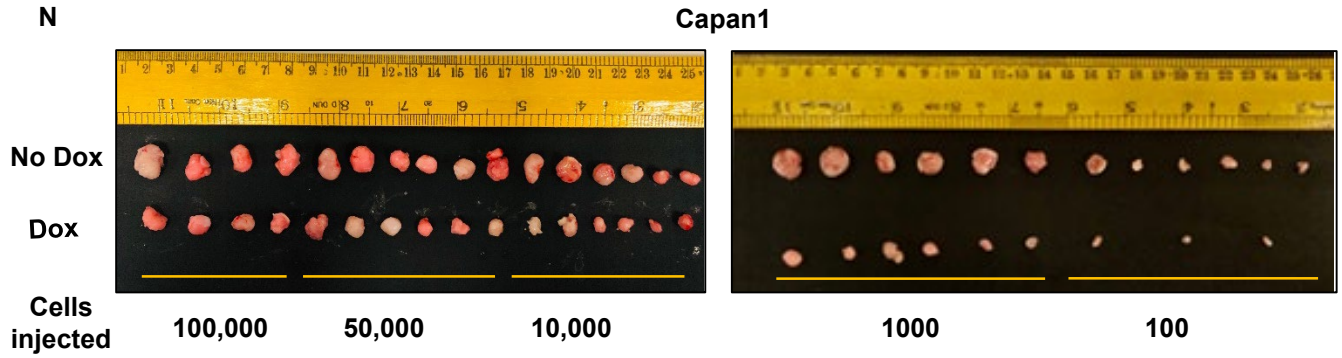
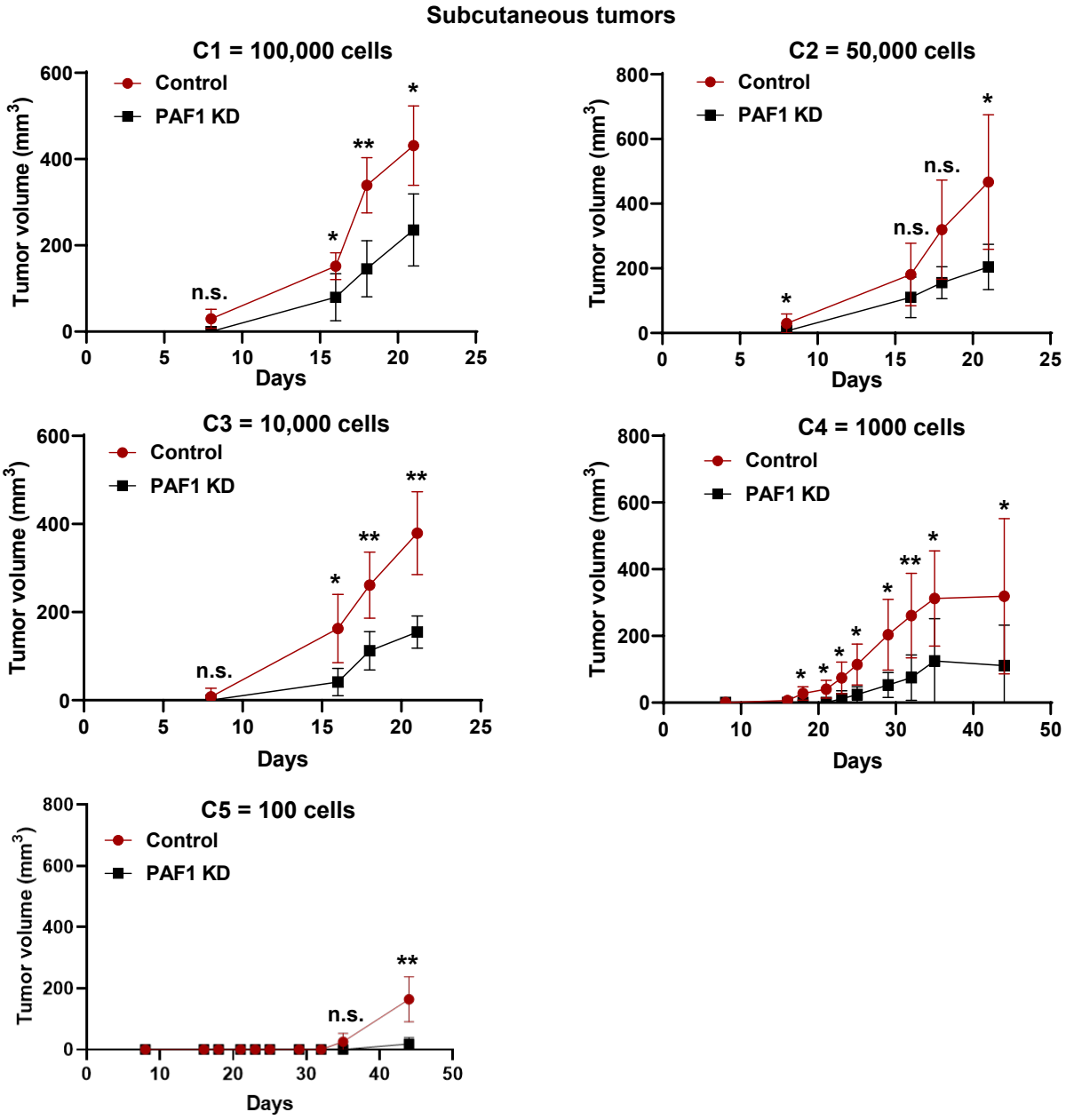

G
SW1990 Control
SW1990 PAF1 KD
Ki 67
H
Control
PAF1 KD
kDa
- - + + Dox
80
PAF1
42
β-ACTIN
I
SW1990
kDa
- + Dox
80
PAF1
34
NANOG
56
SOX9
42
β-ACTIN
J
CD18
kDa
- + Dox
80
PAF1
56
ALDH1
100
CD44v6
56
SOX9
94
β-CATENIN
42
β-ACTIN
K
CD18
*
%age of high CD44+ cells
Control
PAF1 KD
Normalized To Mode
Comp-Red 660_20-A
L
Control
PAF1 KD
DEAB Control
FSC-A
Aldefluor 17.4%
Aldefluor 11.9%
Aldefluor 0.056%
Comp-Blue 530_30-A
CD18
*
% of ALDH+ cells
Control
PAF1 KD
M
Control
PAF1 KD
Verapamil Control
Hoechst Blue-A
Hoechst Red-A
sp 1.03
sp 0.49
sp 0.051
CD18
**
% of SP cells
Control
PAF1 KD

Figure S1


N
Capan1
No Dox
Dox
Cells injected
100,000
50,000
10,000
1000
100


Subcutaneous tumors
C1 = 100,000 cells
C2 = 50,000 cells
C3 = 10,000 cells
C4 = 1000 cells
C5 = 100 cells
Control
PAF1 KD
Tumor volume (mm3)
Days
n.s.
*
**

Supplementary Figure Legends

Figure S1: Loss of PAF1 reduces pancreatic tumorigenesis and proportions of CSCs. A. Immunoblot analysis for PAF1 in CD18 cells representing the efficiency of PAF1 knockdown. **B.** Representative H and E demonstrating histology of pancreas tumor and metastases to liver, intestine, and stomach for control and PAF1-depleted Capan1 xenograft tumors. Scale bars are 1000 μm. **C.** Quantification of the average number of migrating cells in control and PAF1-depleted SW1990 cells. Representative images are shown below. Scale bars are 1000 μm. Data representative of three independent experiments. **D.** Graphical representation of percentage wound closure after 24 h in control and PAF1 depleted SW1990 cells. The representative images are shown on the right. Data representative of three independent experiments and error bars show mean ± S.E. Scale bars are 1000 μm. **E.** Mean tumor weight in control and PAF1 knockdown group generated from orthotopic implantation of CD18 cells in the pancreas of nude mice. Representative image is shown below. Tumors were harvested 6 weeks following implantation. N = 10 mice per group. **F.** Graphical representation of mean tumor weight of control and PAF1-depleted CD18 tumors generated with subcutaneous implantation. Tumors were harvested 6 weeks following implantation. Representative images are shown below. N = 4 mice per group. Both right and left flank were used for implantation. **G.** Representative immunohistochemistry images of Ki67 expression in control and PAF1 depleted SW1990 xenograft tumors. Scale bars are 1000 μm. **H.** Immunoblot analysis of PAF1 from SW1990 control and PAF1 knockdown xenograft tumors. β-ACTIN was used as a loading control. **I.** Immunoblot analysis of Paf1, Nanog, and Sox9 in PAF1 depleted pancreatic tumors obtained from orthotopic implantation of SW1990 cells. **J.** Immunoblot

analysis of PAF1, CSC markers (CD44v6, ALDH1, and Lgr5), and self-renewal markers (β-CATENIN and SOX9) upon PAF1 depletion in CD18 cells. **K.** Percentage of high $CD44^+$ cells in control and PAF1 depleted CD18 cells. Data representative of three independent experiments. Representative histogram from one experiment is shown on the right. **L.** Representative scatter plots and quantification of the number of $ALDH1^+$ cells upon PAF1 depletion in CD18 cells. Data representative of three independent experiments. **M.** Quantification of the percentage of SP cells upon PAF1 knockdown in CD18 cells. Representative scatter plots are on the left. Data representative of three independent experiments. **N.** Capan 1 cells were subcutaneously implanted in left and right flanks of nude mice at different cell concentrations (C1 = 100,000 cells; C2 = 50,000 cells; C3 = 10,000 cells; C4 = 1000 cells; C5 = 100 cells). The top panel represents tumors harvested from control and PAF1 KD cohorts. Tumors generated from cell concentrations C1, C2, and C3 were harvested 3 weeks following implantation. Tumors generated from cell concentrations C4 and C5 were harvested 6 weeks following implantation. N = 6 tumors per cohort per cell concentrations (except C1 that had 4 tumors per cohort as one mouse died prematurely due to unknown reasons). Mean tumor volume for control and PAF1 KD cohort was plotted against the days at which the measurement was made and are depicted as graphs shown below. *P <0.05, **P <0.01, ***P <0.001.

Figure S2

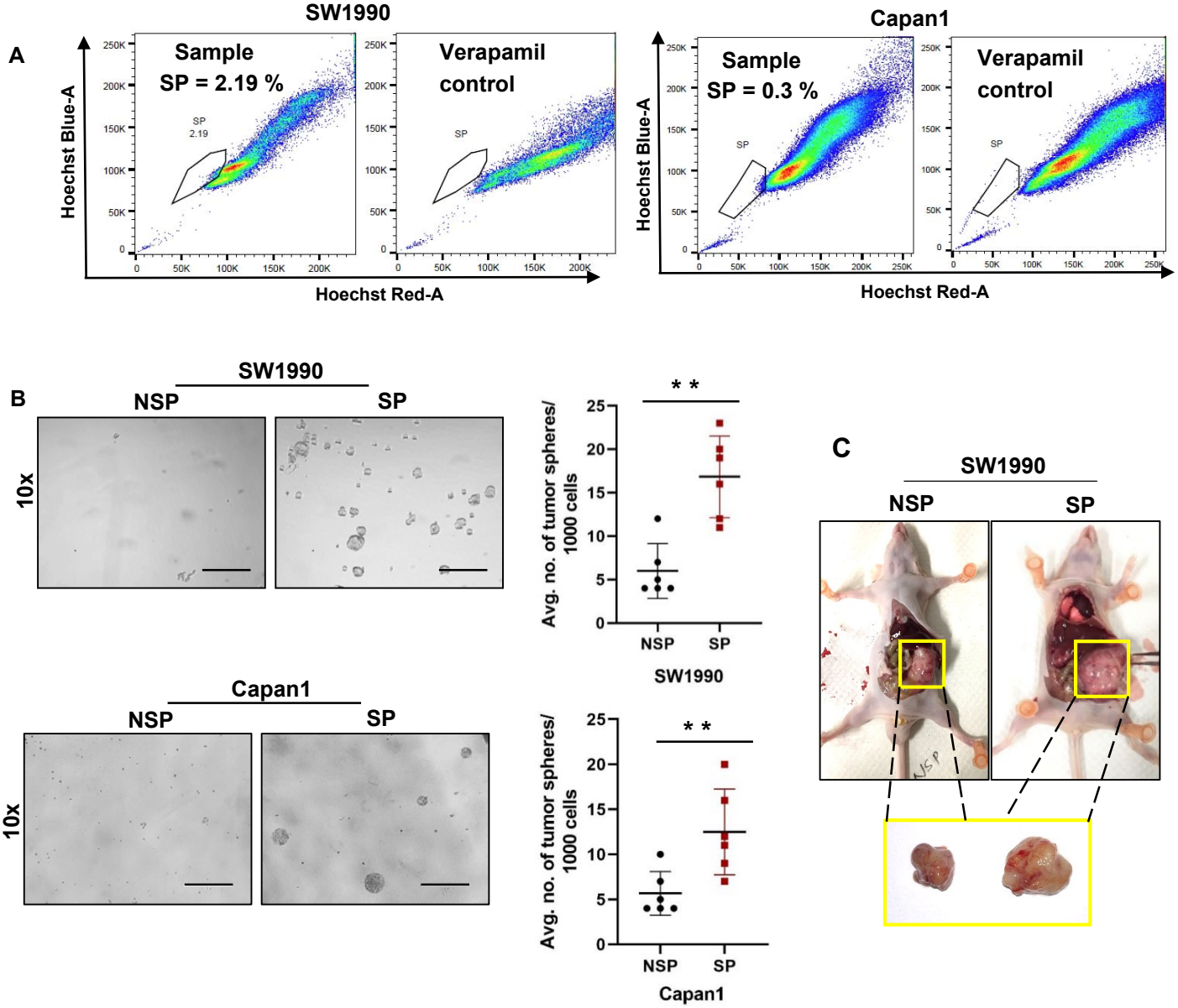

A
SW1990
Sample
SP = 2.19 %
Verapamil
control
Hoechst Blue-A
Hoechst Red-A
Capan1
Sample
SP = 0.3 %
Verapamil
control
B
SW1990
NSP
SP
10x
Avg. no. of tumor spheres/ 1000 cells
SW1990
Capan1
C
SW1990
NSP
SP
D
SW1990 SP
Capan1 SP
PAF1
CD44v6
ALDH1
β-Catenin
Sox9
Fold change relative to actin
Control
PAF1 KD

Figure S2

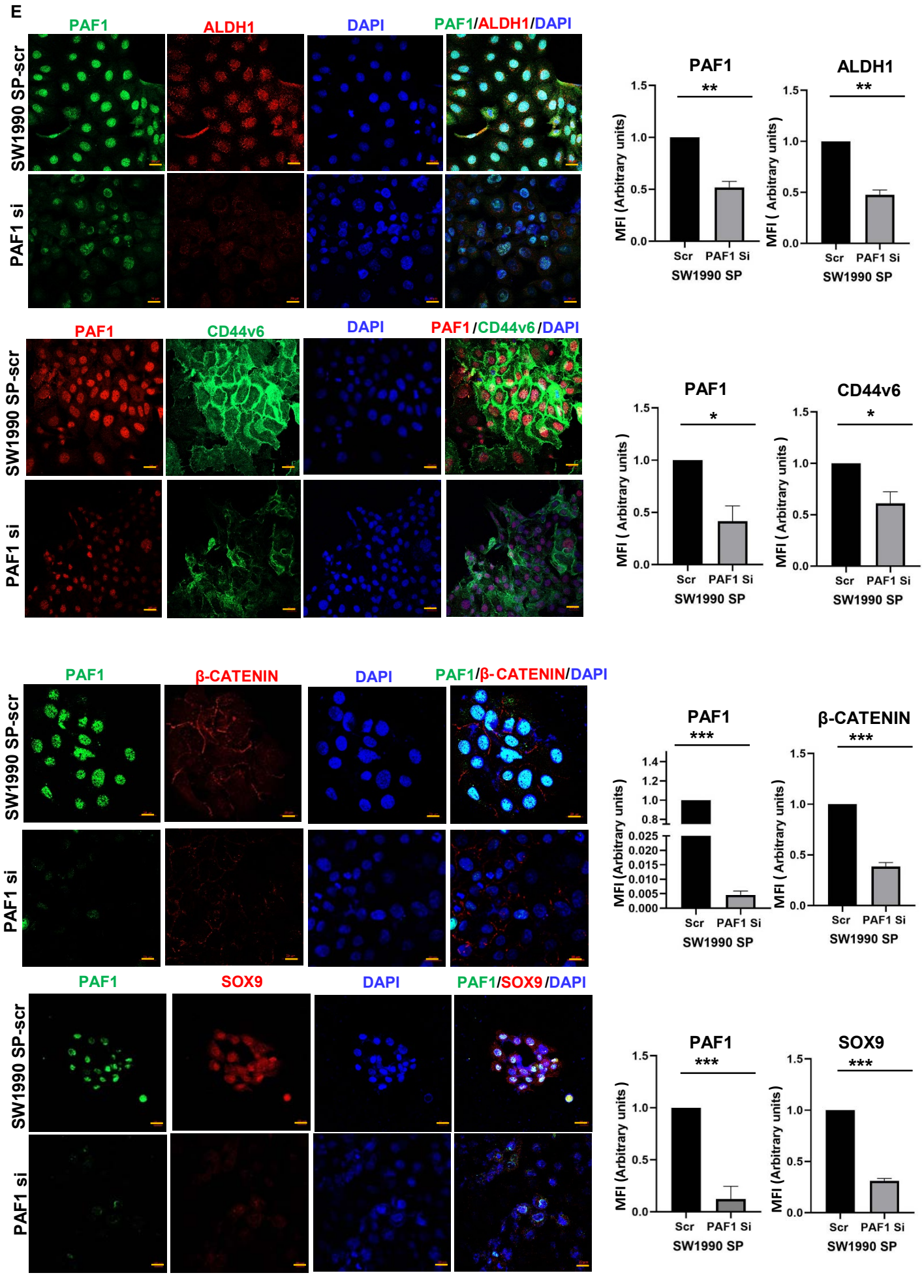

E
PAF1
ALDH1
DAPI
PAF1/ALDH1/DAPI
SW1990 SP-scr
PAF1 si
PAF1
ALDH1
**
MFI (Arbitrary units)
Scr
PAF1 Si
SW1990 SP
PAF1
CD44v6
DAPI
PAF1/CD44v6/DAPI
SW1990 SP-scr
PAF1 si
PAF1
CD44v6
*
MFI (Arbitrary units)
Scr
PAF1 Si
SW1990 SP
PAF1
β-CATENIN
DAPI
PAF1/β-CATENIN/DAPI
SW1990 SP-scr
PAF1 si
PAF1
β-CATENIN
***
MFI (Arbitrary units)
Scr
PAF1 Si
SW1990 SP
PAF1
SOX9
DAPI
PAF1/SOX9/DAPI
SW1990 SP-scr
PAF1 si
PAF1
SOX9
***
MFI (Arbitrary units)
Scr
PAF1 Si
SW1990 SP

Figure S2

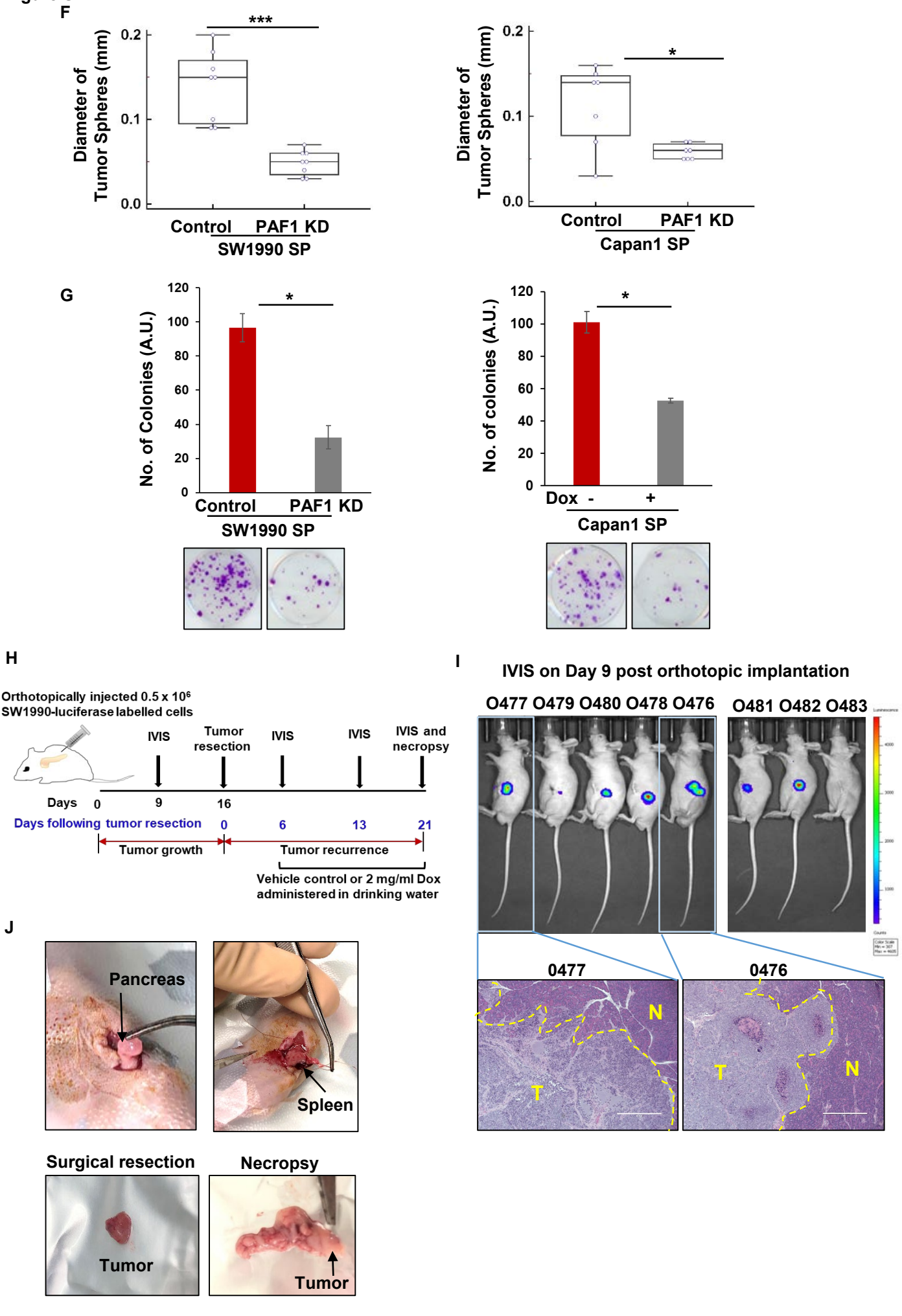
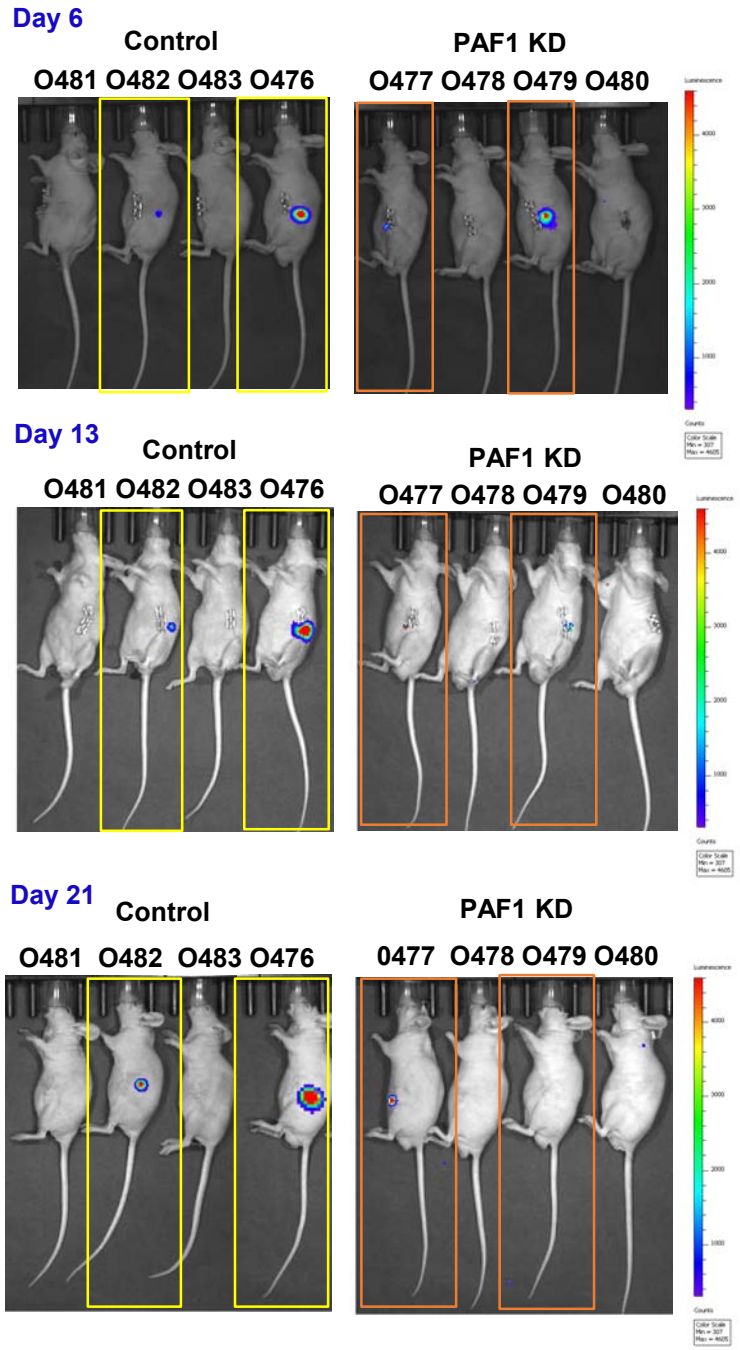
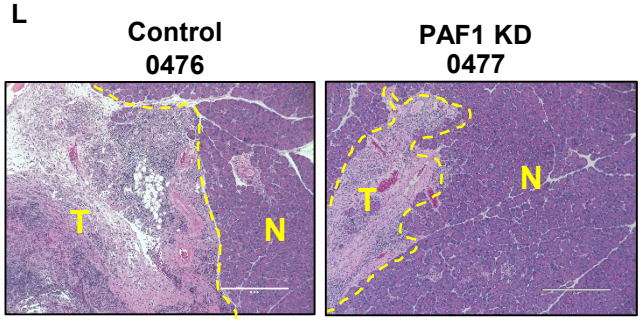

F
Diameter of Tumor Spheres (mm)
0.2
0.1
0.0
***
Control
PAF1 KD
SW1990 SP
Diameter of Tumor Spheres (mm)
0.2
0.1
0.0
*
Control
PAF1 KD
Capan1 SP
G
No. of Colonies (A.U.)
120
100
80
60
40
20
0
*
Control
PAF1 KD
SW1990 SP
No. of colonies (A.U.)
120
100
80
60
40
20
0
*
Dox -
+
Capan1 SP
H
Orthotopically injected 0.5 x 10⁶
SW1990-luciferase labelled cells
IVIS
Tumor resection
IVIS
IVIS
IVIS and necropsy
Days
0
9
16
Days following tumor resection
0
6
13
21
Tumor growth
Tumor recurrence
Vehicle control or 2 mg/ml Dox
administered in drinking water
I
IVIS on Day 9 post orthotopic implantation
O477 O479 O480 O478 O476
O481 O482 O483
0477
0476
N
T
T
N
J
Pancreas
Spleen
Surgical resection
Necropsy
Tumor
Tumor

Figure S2

K


Day 6
Control
O481 O482 O483 O476
PAF1 KD
O477 O478 O479 O480
Day 13
Control
O481 O482 O483 O476
PAF1 KD
O477 O478 O479 O480
Day 21
Control
O481 O482 O483 O476
PAF1 KD
0477 O478 O479 O480

L


Control
0476
PAF1 KD
0477
T
N
T
N

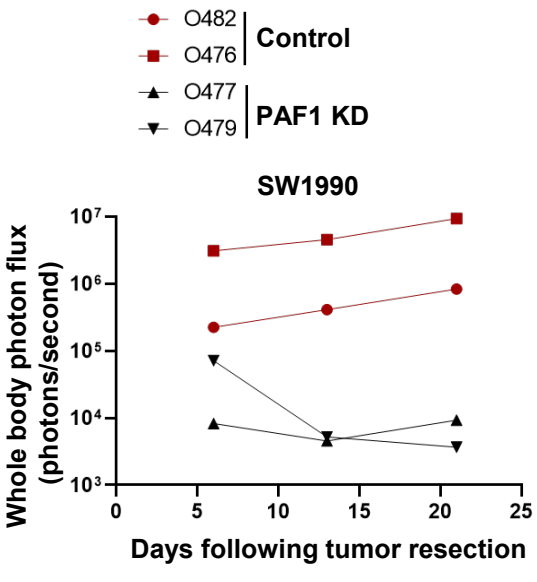

O482
O476
Control
O477
O479
PAF1 KD
SW1990
Whole body photon flux (photons/second)
10^7
10^6
10^5
10^4
10^3
0
5
10
15
20
25
Days following tumor resection

Figure S2: PAF1 is upregulated in pancreatic CSCs and is required for sustenance of CSCs. A. Flow sorting of SP and NSP cells from SW1990 and Capan1 cells. **B.** Representative images of the number and size of tumor spheres formed by SP and NSP cells from SW1990 and Capan1. Graphical representation of the average number of tumor spheres per 1000 cells (mean ± SEM) is shown on the right side. Six replicates per cell line were used. The images were taken two weeks after seeding cells. Scale bars are 1000 μm. **C.** Representative necropsy image of nude mice orthotopically injected with SW1990 SP and NSP cells. 100 cells were injected, and mice were euthanized 4 weeks following implantation. **D.** Quantification of PAF1, CD44v6, ALDH1, β-Catenin, and Sox9 transcripts in control and PAF1 depleted CSCs. Data representative of three independent experiments. Primer sequences are shown in Supplementary table 2. **E.** Immunofluorescence analysis of PAF1, CSC markers (ALDH1 and CD44v6), and self-renewal markers (β-CATENIN and SOX9) in control and PAF1 depleted CSCs. Quantification of MFI is shown on the right. Error bars indicate mean ± S.D. Scale bars are 20 μm. **F.** Quantification of diameter of tumor spheres generated by control CSCs versus PAF1 KD CSCs. **G.** Quantification of total colony area and representative images of colonies formed by control and PAF1 knockdown CSCs. Data representative of three independent experiments and error bars indicate SEM values. **H.** Scheme for surgical tumor resection, IVIS imaging, and necropsy for the assessment of tumor recurrence. **I.** Representative IVIS images from day 9 post orthotopic implantation indicating tumor formation in seven out of eight mice. H and E images depicting histology of xenograft tumors is shown below. The scale bar is 1000 μm. T indicates an area of tumor cells and N indicates an adjacent normal area. **J.** Images showing pancreatic tumor resulting from

orthotopic implantation of PC cells and technique for the surgical resection of the tumor. Lower images depict the resected tumor mass and tumor-bearing pancreas harvested at necropsy. **K.** IVIS images from day 6, 13, and 21 post-surgical resection of xenograft tumors. Yellow boxes indicate mice from the control cohort, while orange boxes highlight the mice that belonged to PAF1 KD cohort. Quantification of photon counts for two mice from each cohort versus days post tumor resection is depicted below. **L.** Representative H and E images of pancreas harvested from control and PAF1 KD cohort at necropsy. Pancreas harvested from control mice showed both areas of cancerous and normal tissue, whereas pancreas from PAF1 KD cohort showed primarily normal areas. Scale bar is 1000 μm. *P <0.05, **P <0.01, ***P <0.001.

Figure S3

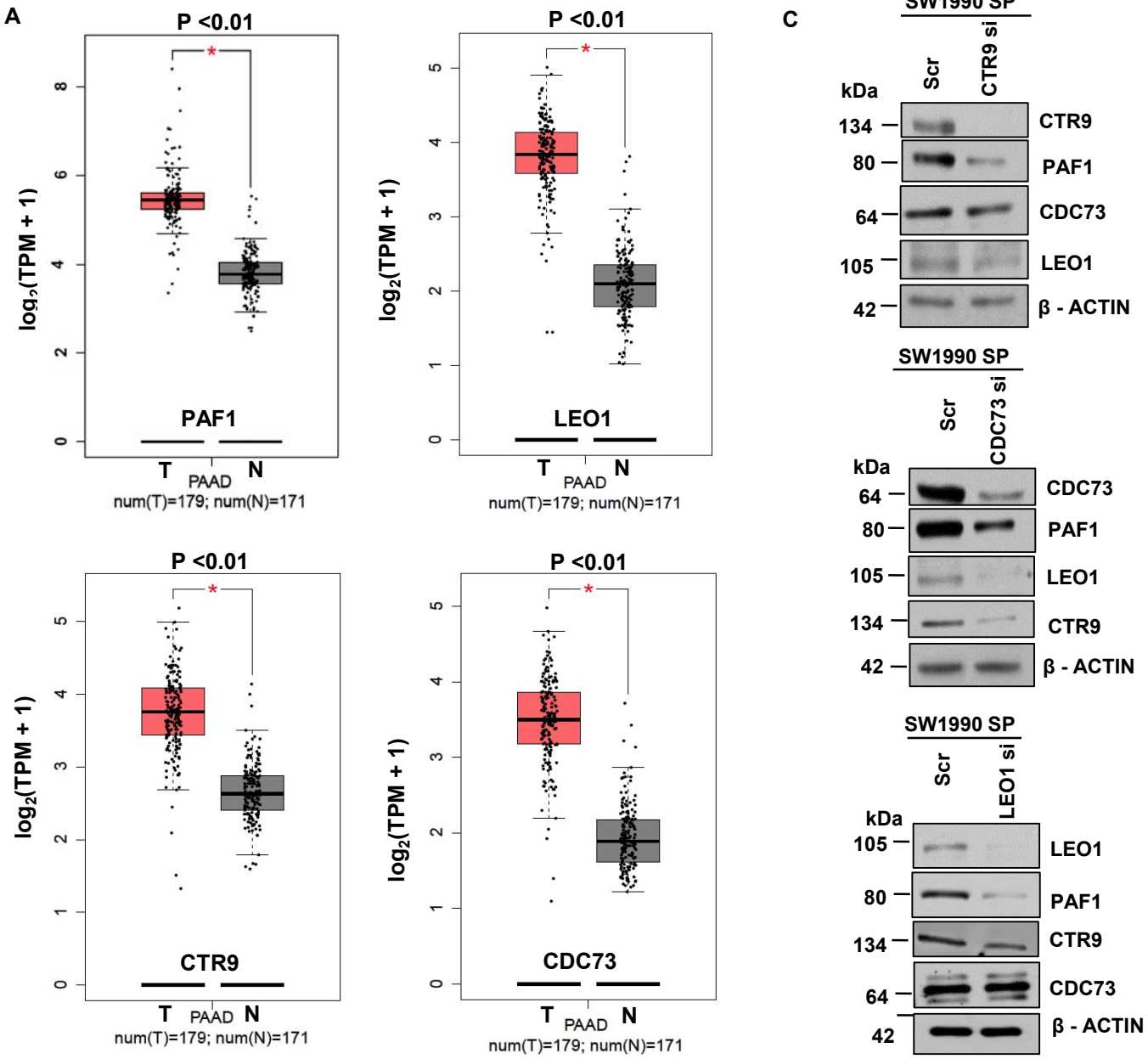
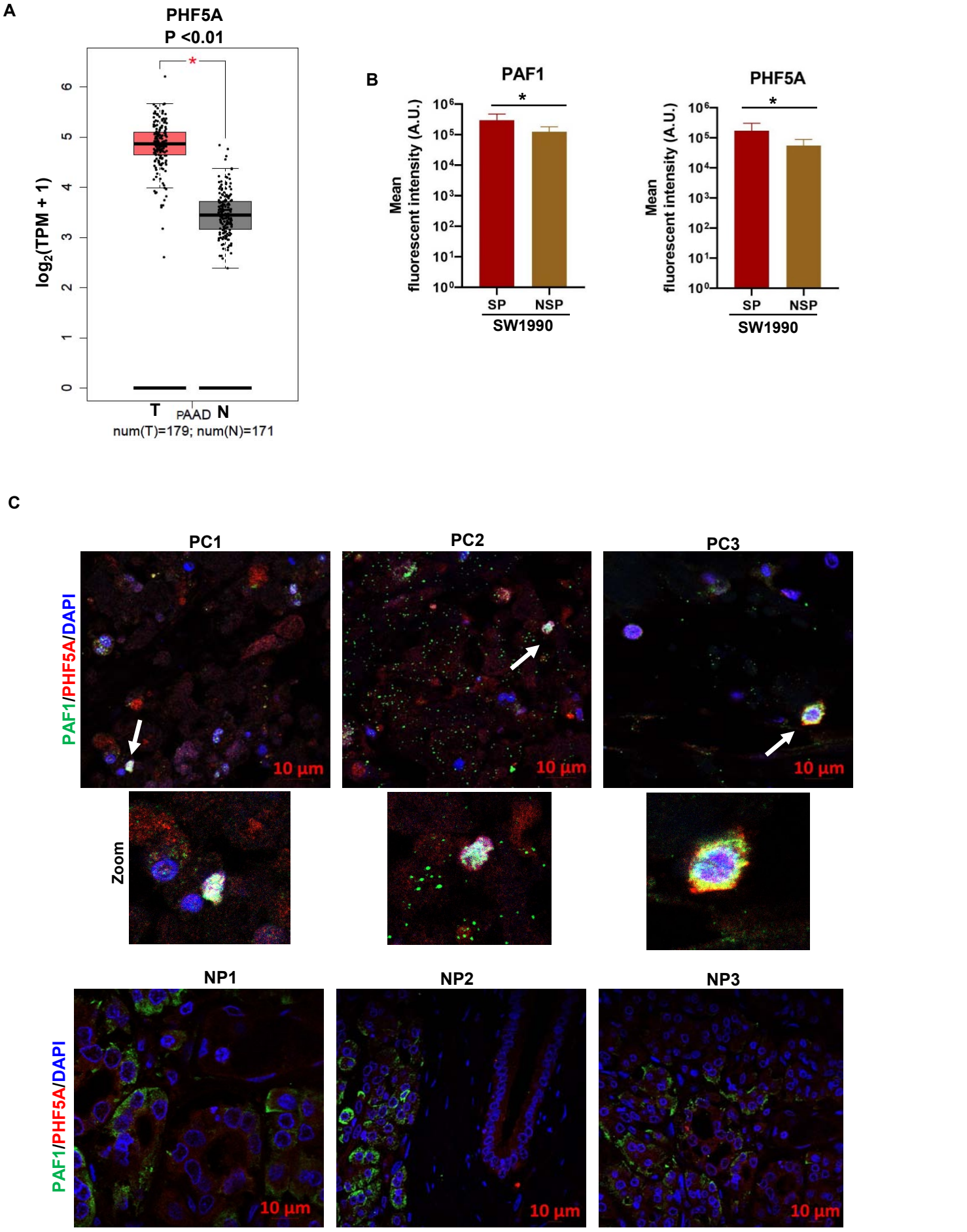

A
P <0.01
log2(TPM + 1)
PAF1
T
N
PAAD
num(T)=179; num(N)=171
P <0.01
log2(TPM + 1)
LEO1
T
N
PAAD
num(T)=179; num(N)=171
P <0.01
log2(TPM + 1)
CTR9
T
PAAD
N
num(T)=179; num(N)=171
P <0.01
log2(TPM + 1)
CDC73
T
PAAD
N
num(T)=179; num(N)=171
C
SW1990 SP
Scr
CTR9 si
kDa
134
CTR9
80
PAF1
64
CDC73
105
LEO1
42
β - ACTIN
SW1990 SP
Scr
CDC73 si
kDa
64
CDC73
80
PAF1
105
LEO1
134
CTR9
42
β - ACTIN
SW1990 SP
Scr
LEO1 si
kDa
105
LEO1
80
PAF1
134
CTR9
64
CDC73
42
β - ACTIN
B
FPKM UQ expression
3e+06
2e+06
1e+06
0e+00
CDC73
CTR9
DDX3X
LEO1
PAF1
PHF5A
SKI8
Gene
Tumor_purity
High
Low
Normal

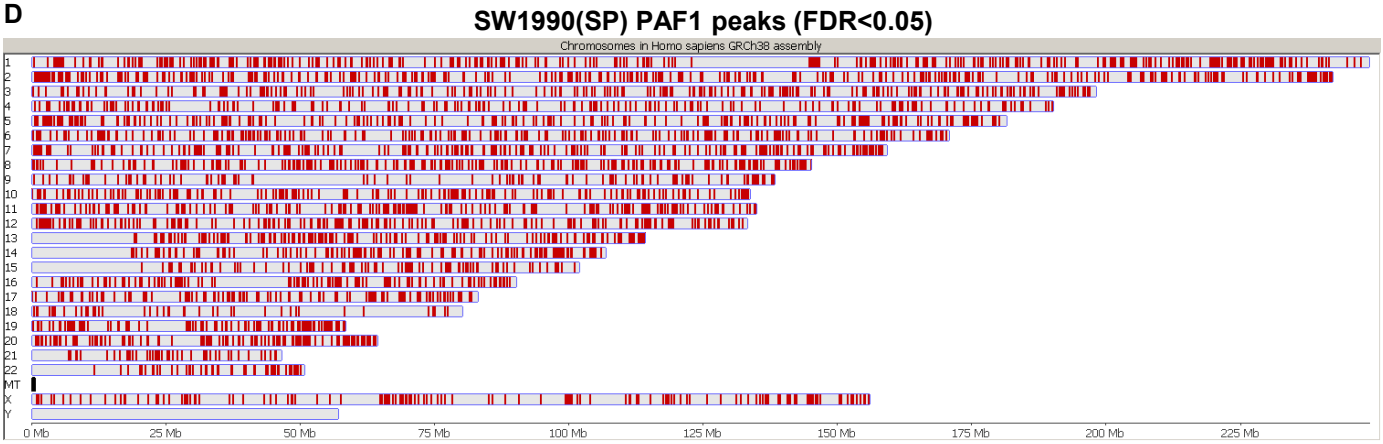
Figure S3: Expression variation of PAF1 and other PAF1C subunits in human PC cases compared to normal pancreas. A. Box plots representing the mRNA expression profiles of PAF1C subunits (PAF1, LEO1, CTR9, and CDC73) across 179 PC cases (TCGA-PAAD data) and 171 normal pancreata (GTEx data) from GEPIA database (here *P <0.01). **B.** Box plot representing the mRNA expression profiles of PAF1C subunits (PAF1, LEO1, CTR9, CDC73, and SKI8), DDX3X, and PHF5A in high tumor purity, low purity samples, and normal pancreas tissues (TCGA-PAAD) obtained from GDC portal. **C.** Immunoblot analysis of PAF1C subunits with individual loss of CTR9, LEO1, and CDC73.

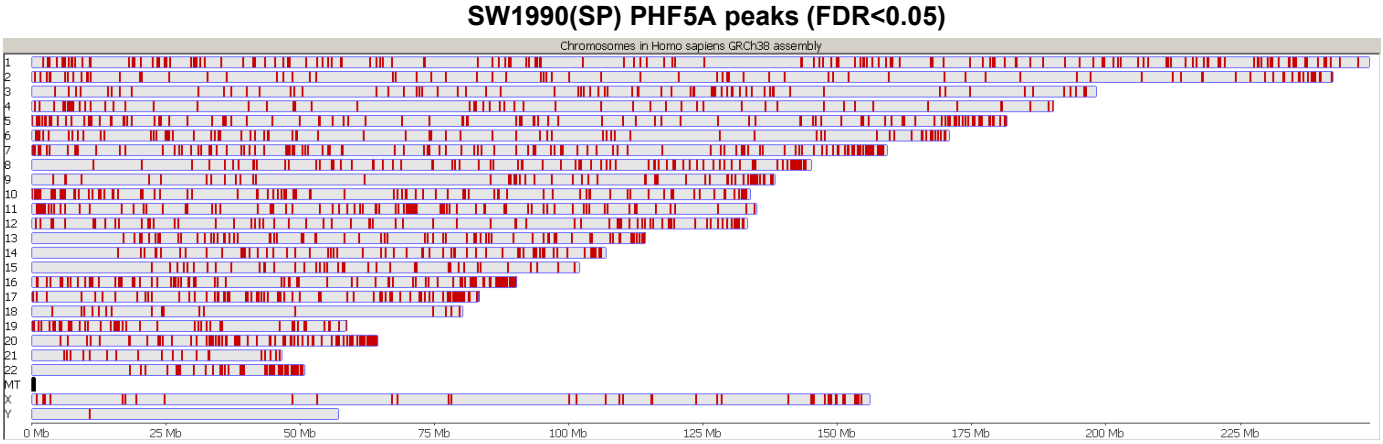
FigureS4


A
PHF5A
P <0.01
log2(TPM + 1)
T
PAAD
N
num(T)=179; num(N)=171
B
PAF1
PHF5A
Mean fluorescent intensity (A.U.)
SP
NSP
SW1990
C
PC1
PC2
PC3
PAF1/PHF5A/DAPI
Zoom
NP1
NP2
NP3
10 μm

FigureS4

D


SW1990(SP) PAF1 peaks (FDR<0.05)
Chromosomes in Homo sapiens GRCh38 assembly
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
MT
X
Y
0 Mb
25 Mb
50 Mb
75 Mb
100 Mb
125 Mb
150 Mb
175 Mb
200 Mb
225 Mb


SW1990(SP) PHF5A peaks (FDR<0.05)
Chromosomes in Homo sapiens GRCh38 assembly
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
MT
X
Y
0 Mb
25 Mb
50 Mb
75 Mb
100 Mb
125 Mb
150 Mb
175 Mb
200 Mb
225 Mb

E

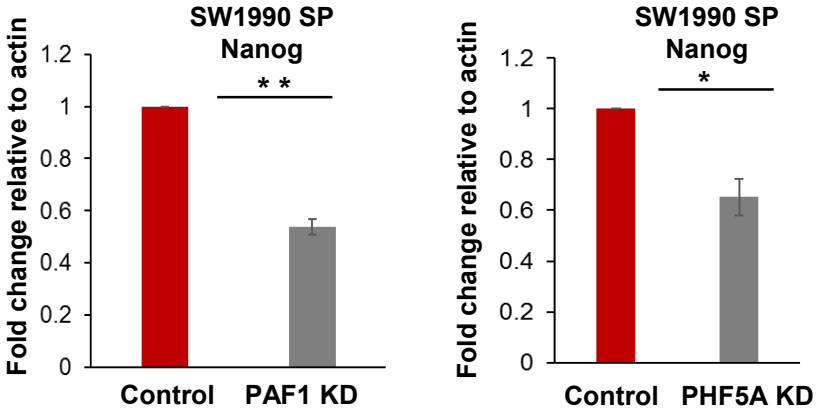
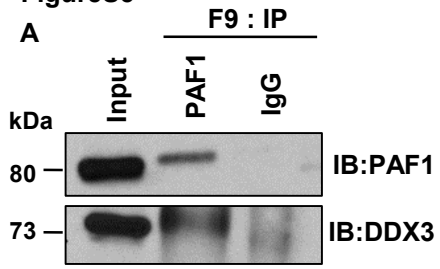
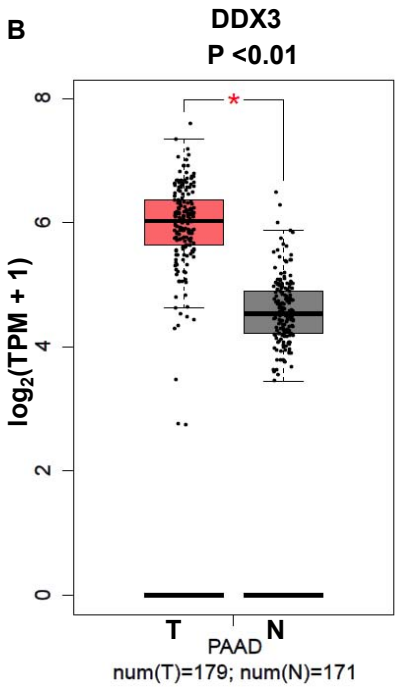
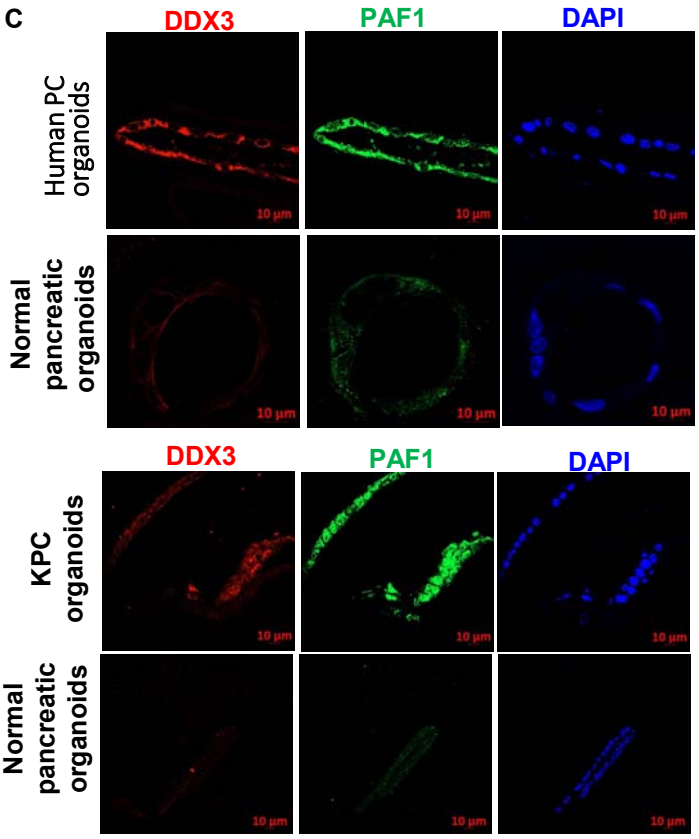

SW1990 SP
Nanog
* *
Fold change relative to actin
1.2
1
0.8
0.6
0.4
0.2
0
Control
PAF1 KD
SW1990 SP
Nanog
*
Fold change relative to actin
1.2
1
0.8
0.6
0.4
0.2
0
Control
PHF5A KD

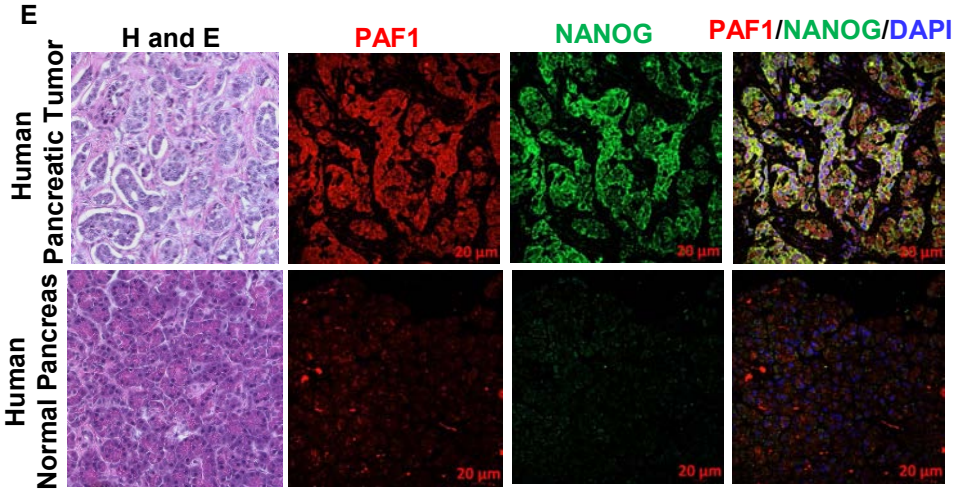
Figure S4: PHF5A is overexpressed in pancreatic CSCs. A. Box plot representing the expression of PHF5A in 171 normal pancreatic tissues (GTEx) and 179 PC (TCGA-PAAD) cases from GEPIA database (here *P <0.01). **B.** Quantification of MFI for PAF1 and PHF5A in SP and NSP cells from SW1990. Data representative of three independent experiments. Error bars indicate mean ± S.E. **C.** Representative immunofluorescence images depicting co-localization of PAF1 with PHF5A in distinct cancer cells within human PC tissues (white arrows). Zoomed images of individual cells showing co-localization are represented in the middle panel. Lower panel shows the expression of PAF1 and PHF5A in normal pancreas. **D.** Chromosome-wide distribution of PAF1 and PHF5A peaks in SW1990 SP cells at a false discovery rate of less than 0.05. **E.** Bar plots representing the expression of Nanog in control versus PAF1 depleted and control versus PHF5A depleted CSCs. Data representative of three independent experiments and error bars indicate mean ± S.E. *P <0.05, **P <0.01, ***P <0.001.

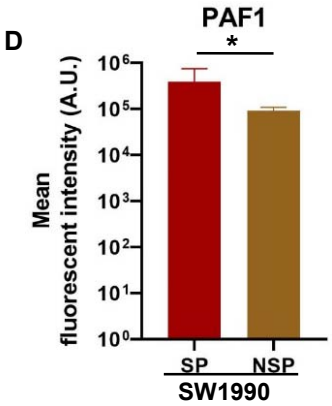
FigureS5

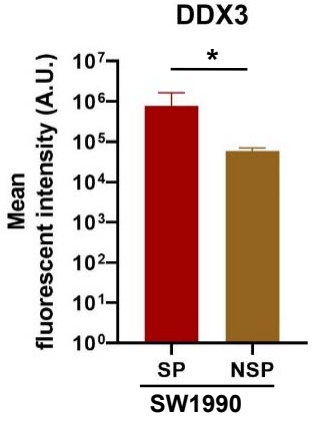

A
F9 : IP
Input
PAF1
IgG
kDa
80
73
IB:PAF1
IB:DDX3


B
DDX3
P <0.01
log2(TPM + 1)
T
N
PAAD
num(T)=179; num(N)=171


C
DDX3
PAF1
DAPI
Human PC organoids
Normal pancreatic organoids
KPC organoids
10 μm


E
H and E
PAF1
NANOG
PAF1/NANOG/DAPI
Human Pancreatic Tumor
Human Normal Pancreas
20 μm


D
PAF1
Mean fluorescent intensity (A.U.)
SP
NSP
SW1990


DDX3
Mean fluorescent intensity (A.U.)
SP
NSP
SW1990

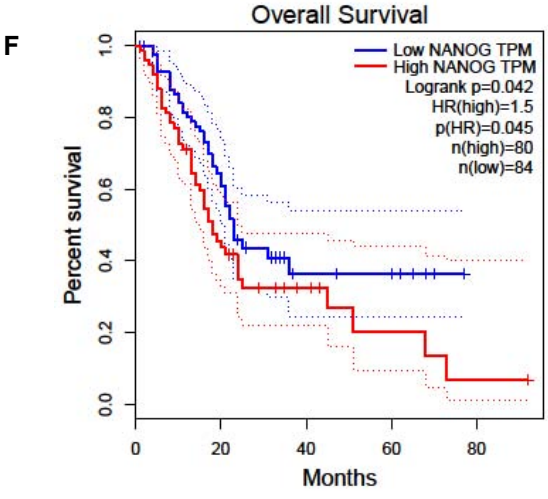

F
Overall Survival
Low NANOG TPM
High NANOG TPM
Logrank p=0.042
HR(high)=1.5
p(HR)=0.045
n(high)=80
n(low)=84
Percent survival
Months

Figure S5: DDX3 is a novel PAF1 interacting partner that is overexpressed in pancreatic CSCs. A. Immunoprecipitation analysis with PAF1 from F9 cells followed by immunoblotting with PAF1 and DDX3. **B.** Box plot representing expression of DDX3 in normal (171 cases) and human PC tissues (179 cases) obtained from GEPIA database (*P <0.01). **C.** Immunofluorescence analysis of PAF1 and DDX3 in human PC and normal pancreatic organoids (top panel), and in mouse KPC and normal pancreatic organoids (lower panel). Data representative of two independent experiments. **D.** Mean fluorescence intensity of PAF1 and DDX3 in SP and NSP cells from SW1990. Data representative of two independent experiments (*P <0.05). **E.** Immunofluorescence analysis of PAF1 and NANOG in human PC tissues and normal pancreas. Tissue histology is represented on the left as Hematoxylin and Eosin stained images. **F.** Kaplan-Meier survival curve for overall survival of PC patients with high and low Nanog expression extracted from GEPIA database.

FigureS6

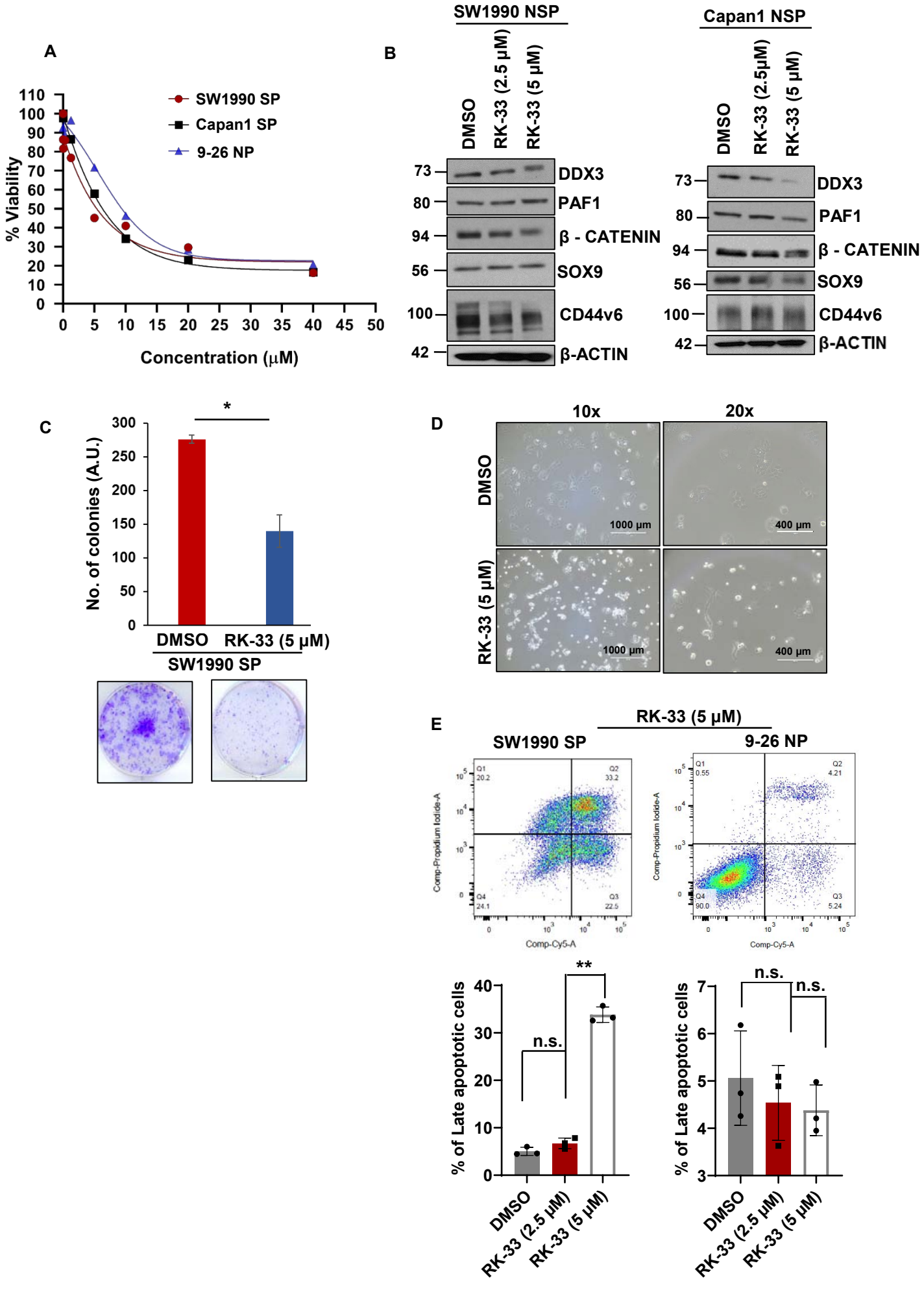
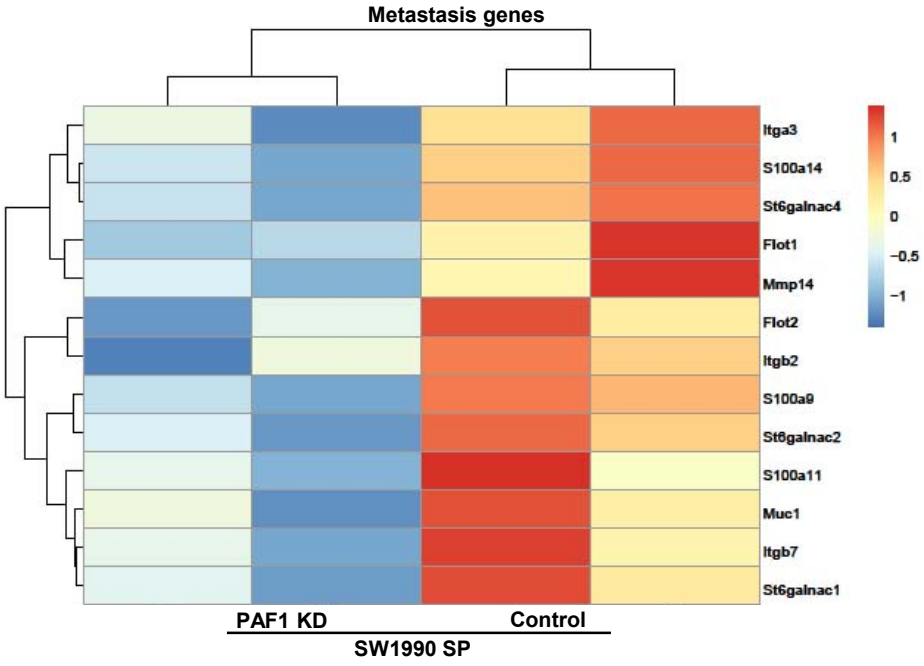

A
% Viability
Concentration (μM)
SW1990 SP
Capan1 SP
9-26 NP
B
SW1990 NSP
Capan1 NSP
DMSO
RK-33 (2.5 μM)
RK-33 (5 μM)
DDX3
PAF1
β - CATENIN
SOX9
CD44v6
β-ACTIN
C
No. of colonies (A.U.)
DMSO
RK-33 (5 μM)
SW1990 SP
D
10x
20x
1000 μm
400 μm
E
RK-33 (5 μM)
SW1990 SP
9-26 NP
Comp-Propidium Iodide-A
Comp-Cy5-A
% of Late apoptotic cells
n.s.

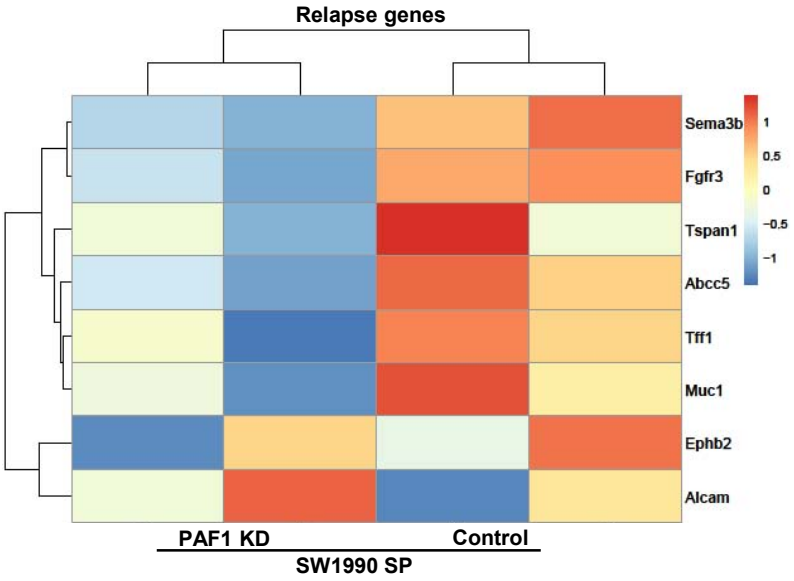
Figure S6: DDX3 inhibitor RK-33 reduces colony formation ability of pancreatic CSCs and induces apoptosis. A. Growth curve representing IC_{50} values for RK-33 in SW1990 SP cells, Capan1 SP cells and 9-26 NP fibroblasts. IC_{50} values are: SW1990 SP = 4.32 μM, Capan1 SP = 5.53 μM, and 9-26 NP = 9.23 μM. SEM values are: SW1990 SP = 0.040, 0.027, 0.042, 0.038, 0.015, 0.022, 0.014, 0.016, 0.014; Capan1 SP = 0.064, 0.036, 0.067, 0.039, 0.069, 0.017, 0.010, 0.010; and 9-26 NP = 0.060, 0.068, 0.077, 0.051, 0.117, 0.062, 0.017, 0.010. **B.** Immunoblotting analysis for DDX3, PAF1, CD44v6, SOX9, and β-CATENIN with RK-33 treatment in SW1990 NSP and Capan1 NSP cells for 48 h. β-ACTIN was used as a loading control. **C.** Graphical representation of total colony area in a colony formation assay with control and RK-33 (5 μM)-treated SW1990 SP cells. Data representative of three independent experiments and error bars indicate SEM vales. Cells were treated for 48 h and colonies were stained 15 days following seeding. **D.** Representative images for morphology of SW1990 SP cells following treatment with RK-33 (5 μM) for 48 h. **E.** Scatter plots demonstrating apoptotic cells following 48 h treatment with RK-33 (5 μM) in SW1990 SP and 9-26 NP cells. Quantification of percentage of late apoptotic (Annexin V and PI double positive) cells is shown below. *P <0.05, **P <0.01, ***P <0.001. n.s. = non-significant.

FigureS7

A



B



C

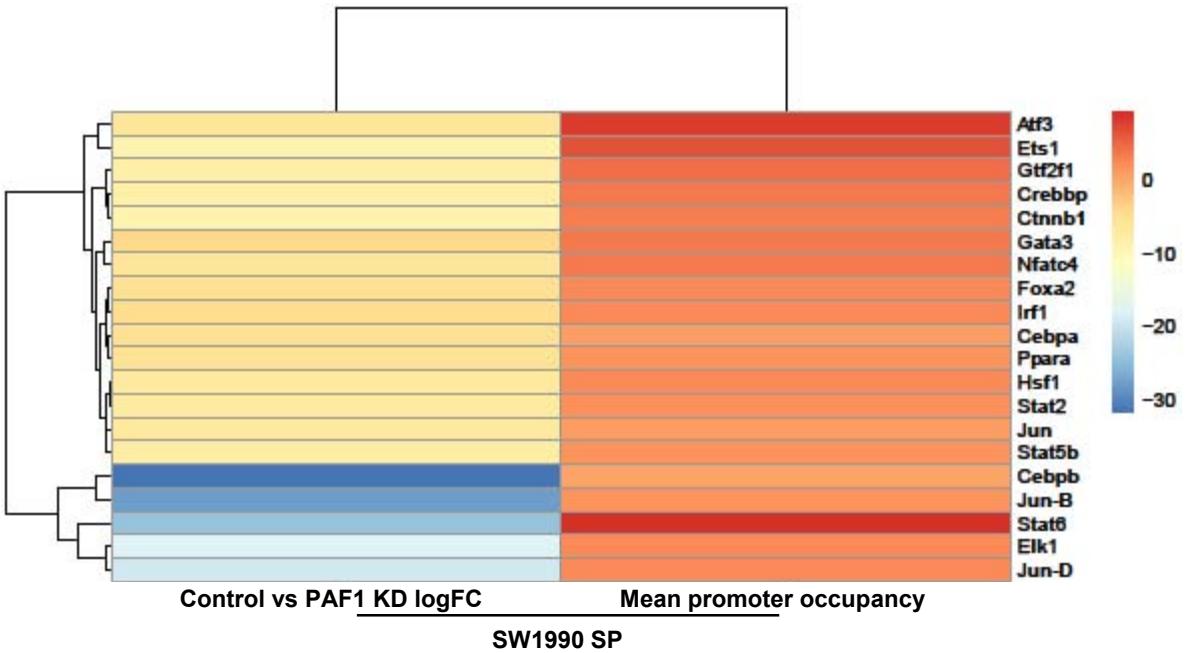


Figure S7: PAF1 regulates an array of stem cell genes. A. Heatmap of differentially expressed metastasis-promoting genes in control and PAF1 depleted SW1990 SP cells. **B.** Heatmap of differentially expressed relapse genes in control and PAF1 depleted SW1990 SP cells. **C.** Heatmap representing differential fold expression of genes downregulated upon PAF1 depletion using PCR array (left panel) and mean promoter occupancy of PAF1 on these gene promoters (right column).

Supplemental Methods

In silico analysis of PAF1C subunits, PHF5A, and DDX3 in PC data sets

The GEPIA web server (http://gepia.cancer-pku.cn/) was used to analyze TCGA PC expression data. Herein, expression data for PAF1, CTR9, LEO1, CDC73, SKI8, PHF5A, and DDX3 were obtained. The data were downloaded using the web portal and subsequently plotted[1]. The normalized PAAD gene expression data (FPKM-UQ) was obtained from Genomic Data Commons (GDC) web portal (portal.gdc.cancer.gov). Grouped boxplots of PAF1, other PAF1C subunits, DDX3, and PHF5A were made using the ggplot2 package in R. The boxplots were grouped based on the ABSOLUTE purity classifications, as provided by Raphael et al., 2017[2].

Orthotopic tumor implantation

SW1990, Capan1, and CD18 cell lines were engineered for inducible PAF1 knockdown using a lentiviral system; these cell lines were injected in the pancreas of 5-week-old athymic nude mice orthotopically as described previously[3]. Doxycycline hyclate (Sigma-Aldrich) was administered to mice assigned to the PAF1 KD group via drinking water at a concentration of 2 mg/ml in 2% sucrose solution. Control mice received 2% sucrose in drinking water.

Subcutaneous tumor implantation

Control and PAF1 KD cells (40,000) were suspended in 500 µl of PBS and mixed with Matrigel (Fisher Scientific) in a 1:1 ratio. Cells (4,000) in 100 µl volume of PBS-matrigel mix were injected subcutaneously into the right and left flanks of 5-week-old athymic nude mice. Four mice were used per group, and the appearance of tumors was checked by palpitation two times per week. Animals were sacrificed soon after the tumor nodule reached 0.8 cm diameter.

Determination of tumor recurrence *in vivo*

Tumor recurrence has previously been evaluated in mouse models of PC[4]. Briefly, 0.5×10^6 luciferase labelled SW1990 cells were orthotopically implanted in 6 week-old nude mice. Tumor formation was ascertained using IVIS. Two weeks post-implantation of cancer cells in the pancreas, tumors were resected as described previously[4]. The primary tumor in the pancreatic

tail was carefully removed by disconnecting the pancreatic body, including blood vessels, by using one clip followed by coagulation and cutting. We confirmed negative margins after the resection to ensure complete tumor removal. Six days after tumor resection, mice were imaged using IVIS to assess tumor burden, and mice were then randomized into two groups: control and PAF1 KD. Mice segregated into PAF1 KD received doxycycline (2 mg/ml) in 2% sucrose in drinking water, and the other group received 2% sucrose only (vehicle control). Tumor-bearing mice in both groups were followed for three weeks and tumor recurrence was evaluated using IVIS and histological analysis at the endpoint.

Limiting dilution assay

Cell suspension of Capan1 cells was prepared for five different concentrations (C1 = 100,000 cells; C2 = 50,000 cells; C3 = 10,000 cells; C4 = 1000 cells; C5 = 100 cells) in a 50 µl volume of PBS. Cells were mixed with 50 µl Matrigel and were subcutaneously implanted into the right and left flanks of 5-week-old athymic nude mice. Mice were segregated into two groups: control and PAF1 KD. Three mice were used per group per cell concentration (6 tumors in each group/cell concentration), and the appearance of tumors was checked by palpitation thrice per week. Tumor dimensions were measured using Vernier Calipers for the determination of tumor volume. Tumors were harvested when control tumors became necrotic.

Organoid generation and culture

In brief, tumor organoids were established after tumor resection, mechanical and enzymatic digestion of pancreatic tumor from KPC autochthonous mouse model with 0.012% (w/v) collagenase XI (Sigma) and 0.012% (w/v) dispase (GIBCO) in DMEM media containing 1% FBS (GIBCO) and embedded in growth factor reduced (GFR) Matrigel (BD)[5]. These organoids were maintained and cultured in complete AdDMEM/F12 medium supplemented with HEPES (Invitrogen), Glutamax (Invitrogen), penicillin/streptomycin (Invitrogen), B27 (Invitrogen), Primocin (1 mg/ml, InvivoGen), N-acetyl-L-cysteine (1 mM, Sigma), mouse recombinant Wnt3a (100 ng/ml, Milipore), human recombinant RSpondin1 (1 µg/ml, Nuvelo), Noggin (0.1 mg/ml,

Peprotech), epidermal growth factor (EGF, 50 ng/ml, Peprotech), Gastrin (10 nM, Sigma), fibroblast growth factor 10 (FGF10, 100 ng/ml, Preprotech), Nicotinamide (10 mM, Sigma), and A83- 01 (0.5 mM, Tocris). Human organoids were derived from the pancreas tissue obtained from fresh tissue bank at UNMC and cultured using the same protocol as followed with mouse organoids with the addition of Wnt3A (1X (50% V/V), Preprotech).

ALDEFLUOR Assay

The cells expressing high levels of activity of the enzyme aldehyde dehydrogenase (ALDH) were stained using ALDEFLUOR reagent (Stem Cell Technologies, Cambridge, MA, USA. Cat: 01700) using the manufacturer's instructions**.** A specific inhibitor of ALDH, diethylaminobenzaldehyde (DEAB), was used to control for background fluorescence.

Analysis of $CD44^+$ CSCs using flow cytometry

Cells were trypsinized, counted, and resuspended to get one million cells/100 µl per tube. Cells were incubated with human CD44-APC antibody (Miltenyi Biotec Inc. Auburn, CA, USA .Cat:130-098-110) for 15 minutes at 4 °C, followed by two PBS washes and resuspension in 500 µl media. Appropriate single and double color controls were included, and cells with a high expression of CD44+ (high CD44+) were gated using appropriate controls.

Apoptosis Assay

SW1990 SP and 9-26 NP cells were treated with vehicle control DMSO or RK-33 (5M) for 48 h. Cell supernatant and adherent cells were collected, stained with annexin V and propidium iodide (PI), and analyzed by flow cytometry as described previously[6].

Immunoblot Assay

Cells were processed for protein isolation and Western Blotting using standard procedures, as described previously [7]. The following primary antibodies were used: anti-PAF1, anti-LEO1, anti-CDC73, anti-CTR9 (Bethyl Laboratories, Montgomery, TX, USA); anti-PHF5A (Proteintech, Rosemont, Illinois, USA); anti-OCT3/4, anti-SOX-2, anti-CD24, anti-ESA (Santa Cruz Biotechnology, Dallas, TX, USA); anti-CD44 (Cell signalling Technology, Danvers, MA, USA);

anti-CD133 (Abnova, Walnut, CA, USA); anti–β-Catenin (Sigma); anti-DDX3, anti-Lgr5, and anti-SOX-9 (Abcam, Cambridge, MA, USA). All antibodies were incubated overnight at 4 °C. β-ACTIN was used as a loading control.

Immunofluorescence analysis

Cells were plated, fixed, and processed as described previously[7]. Primary antibodies specific for mouse-PAF1 (1:100 in PBS), mouse-OCT3/4 (1:100), rabbit-β-CATENIN (1:100), mouse-CD44 (1:250), and rabbit-ESA (1:1500) were used with a 4 h incubation for cells. For human and mouse tissues, we followed the same procedure mentioned previously[8], but incubation with primary antibodies was performed overnight at 4 °C. Following primary antibody incubation, the cells and tissue sections were processed using standard procedures, as described previously[7]. Confocal images were collected using Zeiss LSM800 confocal microscope with a 63×/1.4 NA oil objective.

Quantification of fluorescent intensity: Multichannel snapshots were split into separate channels in ImageJ. A region of interest was drawn around individual cells in one of the two channels using the “freehand” tool. This region was then subjected to the measure plugin, and intensity was calculated by multiplying area and mean intensity. The measurement of intensity was made for 10 cells in a field and 5-6 images per condition. The intensity values were represented relative to the control condition in arbitrary units (A.U.).

Colony formation assay (clonogenic assay)

SW1990 SP PD2/PAF1 CRISPR KD cells and control SW1990 SP cells were trypsinized and seeded at a density of 1000 cells/well in a 6-well plate in triplicates. The cells were cultured in CSC-specific media, with media changed once in two days. After two weeks of growth, the cells were fixed with 100% methanol and stained with crystal violet stain (0.1%, w/v in 20 nm 4-morpholinepropanesulfonic acid; Sigma) before the colonies started to merge. Total colony area was quantified using ImageJ.

Tumor sphere assay

Capan1 SP cells that were engineered to knockdown PAF1 using Dox-induced shRNA were seeded in triplicates in a 96-well non-adherent plate (Corning Inc., Corning, New York, USA) in CSC-specific media at a concentration of 100 cells/well in two sets. The cells in suspension culture were observed under the microscope and fresh media was added every alternate day without removing the existing media in the set designated as control, whereas the set designated as KD set received Dox at 2 μg/ml every alternate day. Treatment with RK-33 was initiated 24 h after cell seeding. A week later, multiple images were taken per well for different fields of view. The diameter of each tumor sphere was measured using Motic Images Plus 2.0 ML software; the plot depicting the diameter of tumor sphere in PAF1 KD cells and control cells was plotted using MedCalc software.

Cell motility assay

The motility assay was performed by using a chamber containing a monolayer-coated polyethylene teraphthalate membrane (six-well insert, a pore size of 8 μm; Becton Dickinson, Franklin Lakes, NJ, USA). Both control and PAF1 knockdown (Dox treated) SW1990 cells (1×10^6 cells per well) were seeded in six-well plates. After a 24 h incubation period, the migrated cells that had reached the lower chamber were stained with a Diff-Quick stain set and counted in different fields. The average number of migrated cells per representative field was calculated.

Wound-healing assay

The wound-healing assay was performed as described previously[9]. Briefly, 4×10^6 control and PAF1 KD (2 mg/ml doxycycline treated) cells were seeded in 6-well plates. After 24 h, a scratch was made on the bottom of each well using a P200 pipette tip. Photographs of the scratch were taken at 0 h and 24 h. The pictures were then analyzed to quantify the change in size of the scratches as the cells migrated into the space; wound closure was compared between control and PAF1 depleted cells.

Immunoprecipitation analysis

SW1990 SP, HPDE, and F9 cells were cultured in their respective culture conditions, and the lysates were collected in a CHAPS non-denaturing immunoprecipitation buffer (20 mM Tris, pH 7.5, 200 mM NaCl, 1%). Immunoprecipitation was performed with PAF1 (ChIP-grade, Abcam) and anti-Phf5a (Proteintech) antibodies using Dynabeads protein G (Invitrogen). Following pre-clearing, 500 μg of lysates prepared at 1 mg/ml were incubated with 3 μg of respective antibodies overnight in a rotor at 4 °C. The next day, dyna beads (20 μl per reaction) were added to the antibody-lysate mix and incubated for rotation at 4 °C for 6 h. Thereafter, the unbound antibodies were washed with IP wash buffer (40 mM HEPES pH 7.4, 500 mM NaCl, 2.5 mM $MgCl_2$, 2 mM EGTA, and 1% Triton X-100), and the immunoprecipitated proteins were eluted using a high pH ammonium buffer and were concentrated using a speed Vac. The immunoprecipitates, or total cell lysates, were transferred onto the PVDF membrane after being resolved on 10% SDS PAGE, and thereafter were incubated overnight at 4 °C with primary antibodies (anti-PAF1, anti-OCT3/4, anti-LEO1, anti-CTR9, and anti-CDC73).

RNA isolation and quantitative reverse-transcription Polymerase Chain Reaction (qRTPCR)

Total cellular RNA was extracted from cells using the RNAeasy kit (Qiagen, Hilden, Germany) and processed for reverse transcription to generate cDNA as described previously[10]. cDNA products were assayed by quantitative real-time PCR using SYBR Green incorporation. The expression of all genes was normalized to that of internal control β-actin and expressed relative to the indicated reference sample (average ± SEM of triplicate reactions).

Human transcription factors PCR array

Total RNA isolated from SW1990 SP and PAF1-depleted SP cells were reverse-transcribed using RT^2 SYBR qPCR master mix (330,401, Qiagen). Aliquots of 25 μl of the mix of both samples were added in separate 96-well PCR array kits containing lyophilized gene-specific

primer set (PAHS-075Z, Qiagen). Threshold cycles were used to calculate fold change using the online web server of RT^2 profiler PCR array data analysis[11].

Chromatin Immunoprecipitation (ChIP)-Sequencing

SW1990 SP cells were fixed with 0.4% formaldehyde and 1.5 mM EGS (ethylene glycol bis(succinimidyl succinate)) (Thermo Scientific Fischer), washed, harvested, and resuspended in 500 µl SDS lysis buffer (1% SDS, 10 mM EDTA, 50 mM Tris-HCl [pH 8.1], 1 mM PMSF, and 1 µg/ml aprotinin). Samples were sonicated and diluted in ChIP dilution buffer (0.01% SDS, 1.1% Triton X-100, 1.2 mM EDTA, 16.7 mM Tris-HCl [pH 8.1], 167 mM NaCl, 1 mM PMSF, and 1 µg/ml aprotinin). For input control, 10% of sonicated samples were separated. Immunoprecipitation was performed with ChIP-grade PAF1 (Abcam) and Phf5a (Proteintech) antibody as previously described. Chromatin extracts were pulled down with protein A+G beads. The samples were washed extensively with wash buffers (low salt, high salt, LiCl, and Tris/EDTA buffers), eluted with SDS elution buffer, and subjected to reverse cross-linking and proteinase digestion. IgG antibody was used as a control for the ChIP assay.

ChIP-seq Analysis: The captured and purified DNA was prepared for high throughput sequencing using the New England Biolabs NEBNext Ultra II DNA Library Prep Kit for Illumina. The resulting indexed libraries were sequenced by the UNMC Sequencing Core Facility using an Illumina NextSeq 500 Genome Analyzer. Initial raw sequence files were processed based on the following steps. Adaptor sequences and low quality (Phred score < 20) ends were trimmed from sequences using Trim Galore software package (http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/). Resulting fastq files were aligned to the human genome (GRCh38/hg38) using the sequence aligner Bowtie2 (version 2.2.3)[12]. The software package Picard routine, MarkDuplicates, (http://broadinstitute.github.io/picard/) was used to remove sequence duplications. For peak calling of ChIP enriched regions, the MACS2 peak caller software (version 2.1.1)[13] of each ChIP

to corresponding input DNA sample was used to determine binding regions based on a FDR adjusted p-value (q-value) <0.05. BigWig files were generated using the deeptools bamCoverage routine (https://deeptools.readthedocs.io/en/develop/). Alignment of significant peaks to gene-specific regions was accomplished using the bedtools routine intersect (https://bedtools.readthedocs.io/en/latest/). Raw files were submitted to GEO (accession number GSE144371).

RNA-Sequencing

SW1990 SP control and PAF1 CRISPR KD cells were lysed in RNA lysis buffer supplied with the mirVana RNA isolation kit (Thermo Scientific Fischer, n = 2). Total RNA was isolated in accordance with the manufacturer's protocol. Whole transcriptome analysis (RNA sequencing) was performed on PAF1 KD and control cells at the Sequencing Core Facility at the University of Nebraska Medical Center. RNA quality was checked with an Agilent Bioanalyzer (Agilent Technologies, Inc.) and all RINs (RNA integrity numbers) were 10. Library preparation was achieved using Illumina TruSeq RNA Sample Preparation Kit according to the manufacturer's protocol. Fifteen cycles of PCR amplification were performed for each library, followed by examination of size distribution using an Agilent Bioanalyzer (Agilent Technologies, Inc.) and a DNA 1000 chip. Each sample used one out of twelve unique indices (Illumina). All libraries displayed a band between 200–500 bp with a peak at approximately 290 bp. The Qubit 2.0 Fluorometer (Life Technologies, Inc.) was used to quantitate the libraries. Loading was performed at a concentration of 6 pM. Sequencing was done on HiSeq 2500 sequencer in rapid mode. It was a single read, 50 cycle, sequencing run, and onboard clustering and V2 chemistry was used. Raw files for RNA-seq were submitted to GEO (accession number GSE144371).

Mass Spectrometry

Sample prep summary: Protein fractions were excised from SDS-PAGE gel, destained, reduced with tris-carboxyethylphosphine, alkylated with iodoacetamide, and digested overnight with

sequencing-grade trypsin (Promega). Tryptic peptides were eluted from the gel and concentrated to 20 µl by vacuum centrifugation and analyzed using a high-resolution mass spectrometry nano-LC-MS/MS Tribrid system, Orbitrap Fusion™ Lumos™ coupled with UltiMate 3000 HPLC system (Thermo Scientific). 500 ng of peptides were run by the pre-column (Acclaim PepMap™ 100, 75µm × 2cm, nanoViper, Thermo Scientific) and the analytical column (Acclaim PepMap™ RSCL, 75 µm × 50 cm, nanoViper, Thermo Scientific). The samples were eluted using a 120-min linear gradient of ACN (5-45%) in 0.1% FA.

Results summary: All MS/MS samples were analyzed using Mascot (Matrix Sciences, London, UK, version 2.6.). Mascot was set up to search the SwissProt database (selected for Homo sapiens, 2018_06, 20361 entries), assuming the digestion enzyme trypsin. Parameters on MASCOT were set as follows: Enzyme: trypsin, Max missed cleavage: 1, Peptide charge: 1+, 2+, and 3+, Peptide tolerance: ± 0.8 Da, Fixed modifications: carbamidomethyl (C), Variable modifications: oxidation (M), MS/MS tolerance: ± 0.6 Da with error tolerant search, Instrument: ESI-TRAP. MASCOT results for different gel cuts of the same sample were combined and analyzed using Scaffold, which allows multiple search results to be condensed into a single result file.

Scaffold (version 4.8.7, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 95.0% probability by the Peptide Prophet algorithm[14] with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least 2 identified peptides. Protein probabilities were assigned by the Protein Prophet algorithm[15]. Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony. Proteins sharing significant peptide evidence were grouped into clusters**.**

Statistical analysis

Student's *t*-test was used to determine the statistical significance between control and PAF1 knockdown group in all the experiments pertaining to this study. Statistical analysis and the generation of graphs were performed using GraphPad Prism and MedCalc software. Alfa level of 0.05 was used for statistical significance unless otherwise indicated. Error bars were given on the basis of calculated standard error values.

Supplementary references

1. Tang Z, Li C, Kang B, et al. GEPIA: a web server for cancer and normal gene expression profiling and interactive analyses. Nucleic Acids Res 2017;45:W98-W102.
2. Cancer Genome Atlas Research Network. Electronic address aadhe, Cancer Genome Atlas Research N. Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma. Cancer Cell 2017;32:185-203 e13.
3. Vaz AP, Deb S, Rachagani S, et al. Overexpression of PD2 leads to increased tumorigenicity and metastasis in pancreatic ductal adenocarcinoma. Oncotarget 2015.
4. Gürlevik E, Fleischmann-Mundt B, Brooks J, et al. Administration of Gemcitabine After Pancreatic Tumor Resection in Mice Induces an Antitumor Immune Response Mediated by Natural Killer Cells. Gastroenterology 2016;151:338-350.e7.
5. Boj SF, Hwang CI, Baker LA, et al. Organoid models of human and mouse ductal pancreatic cancer. Cell 2015;160:324-38.
6. Lakshmanan I, Batra SK. Protocol for Apoptosis Assay by Flow Cytometry Using Annexin V Staining Method. Bio Protoc 2013;3.
7. Vaz AP, Ponnusamy MP, Rachagani S, et al. Novel role of pancreatic differentiation 2 in facilitating self-renewal and drug resistance of pancreatic cancer stem cells. Br. J Cancer 2014;111:486-496.
8. Nimmakayala RK, Seshacharyulu P, Lakshmanan I, et al. Cigarette Smoke Induces Stem Cell Features of Pancreatic Cancer Cells via PAF1. Gastroenterology 2018;155:892-908 e6.
9. Lakshmanan I, Rachagani S, Hauke R, et al. MUC5AC interactions with integrin beta4 enhances the migration of lung cancer cells through FAK signaling. Oncogene 2016;35:4112-21.
10. Ponnusamy MP, Deb S, Dey P, et al. RNA polymerase II associated factor 1/PD2 maintains self-renewal by its interaction with Oct3/4 in mouse embryonic stem cells. Stem Cells 2009;27:3001-3011.
11. Kumar S, Das S, Rachagani S, et al. NCOA3-mediated upregulation of mucin expression via transcriptional and post-translational changes during the development of pancreatic cancer. Oncogene 2015;34:4879-89.
12. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357-9.
13. Zhang Y, Liu T, Meyer CA, et al. Model-based analysis of ChIP-Seq (MACS). Genome Biol 2008;9:R137.
14. Keller A, Nesvizhskii AI, Kolker E, et al. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 2002;74:5383-92.
15. Nesvizhskii AI, Keller A, Kolker E, et al. A statistical model for identifying proteins by tandem mass spectrometry. Anal Chem 2003;75:4646-58.

Supplementary Table 1: Primer sequence for binding sites (B.S.) on Nanog promoter

Name	Forward Primer Sequence	Reverse Primer Sequence	Amplicon Size (bp)
B.S.1	CATTTGGCATGTGTGTCAACTC	TTCAAGTGATAGGATTTGGATAGGG	92
B.S.2	GGATTTGGTCAGCTCCTTTACT	GAACCAGAACGACTCCATCTTC	181
B.S.3	GTCTCAGCCTCCCTAGTAGAT	CACCTGTAATCCCAGCACTT	174
B.S.4	TCAGCCTCGGTGAGTCTTGGTG	CTCCCACACAAGCTGACTTT	310
B.S.5	CATAATCGGGATTTGCTAAGAGTTT	TGTGGGTGTGTGTGTTTCT	211
B.S.6	TGCCTTGGCTTCATGCTATAA	CTGAGGTTATTGAAATTCTCATTAGGG	168
B.S.7	GCCTCCCAATTTACTGGGATTA	CGAGCAACAGAACCTGAAGA	178
B.S.8	TGAGACTGGTAGACGGGATTA	GAAATAGGACCTCCAGAAGGAAA	222

Supplementary Table 2: Primers used in q-RTPCR

Gene Name	(Forward and Reverse primers)	Amplicon Size (bp)
PAF1/PD2	F: CTCACAGCATTACAGCAAACC	231
	R: GTCTCTTCTACAGGCAGGAAAT	
Sox9	F: GAGCCGGATCTGAAGAGGGA	151
	R: GCTTGACGTGTGGCTTGTTC	
CD44v6	F: CCAGGCAACTCCTAGTAGTACAACG	112
	R: CGAATGGGAGTCTTCTTTGGGT	
β-Catenin	F: AAAATGGCAGTGCGTTTAG	100
	R: TTTGAAGGCAGTCTGTCGTA	
ALDH1A3	F: ATCAACTGCTACAACGCCCT	98
	R: TATTCGGCCAAAGCGTATTC	
CTR9	F: AAATTCTCGGCTCTCTCTATGC	92
	R: GGGATACTGTTCTGTGACCTTC	
LEO1	F: AGTGTAGAGCCCAGACCTT	143
	R: CCTTCTTCATCTCGGCGTATC	
CDC73	F: GCCTTCCATCTGAAGTATGATGA	142
	R: TGTACCTGTCCAATGTTTCCC	
β-actin	F: GGACATCCGCAAAGACCTGTA	143
	R: GCTCAGGAGGAGCAATGATCT	
Nanog	F: CCTGTGATTTGTGGGCCTG	78
	R: GACAGTCTCCGTGTGAGGCAT	

List of antibodies

ANTIBODIES	SOURCE	IDENTIFIER AND DILUTION
List of primary antibodies used in Western blot		
Non phospho β-CATENIN	Cell Signaling Technology	Cat# CST-19807P (Rb) (1:1000)
ALDH1	Santa Cruz Biotechnology	Cat# sc-374149 (Ms) (1:1000)
CD44	Cell Signaling Technology	Cat# CST-5640S (Ms) (1:1000)
OCT3/4	Santa Cruz Biotechnology	Cat# sc-5279 (Ms) (1:500)
PAF1/PD2	Bethyl Laboratories	Cat# Bethyl-A300-173A (Rb) (1:5000)
CD44v6	Thermo Fisher Scientific	Cat# BMS125 (Ms) (1:1000)
CD24	Santa Cruz Biotechnology	Cat# sc-19585 (Ms) (1:1000)
ESA (EpCAM)	Cell Signaling Technology	Cat# CST 2626 (Rb) (1:1000)
CD133	Cell Signaling Technology	Cat# CST-5860 (Rb) (1:1000)
DDX3	Abcam	Cat# ab235940 (Rb) (1:1000)
LEO1	Bethyl Laboratories	Cat# A300-174A (Rb) (1:1000)
CTR9	Bethyl Laboratories	Cat# A301-395A-1 (Rb) (1:1000)
CDC73	Bethyl Laboratories	Cat# A300-170A (Rb) (1:1000)
β-Actin	Santa Cruz Biotechnology	Cat# sc-47778 (Ms) (1:2000)
PHF5A	Proteintech	Cat# 15554-1-AP (Rb) (1:500)
SOX2	Santa Cruz Biotechnology	Cat# sc-20088 (Rb) (1:500)
SOX9	Abcam	Cat# ab26414 (Rb) (1:1000)
NANOG	Thermo Fisher Scientific	Cat# PA1-097 (Rb) (1:1000)
List of secondary antibodies used in Western blot		
Goat anti-Rb IgG HRP	Invitrogen	Cat# 31460 (1:3000 to 1:5000)
Goat anti-Ms IgG HRP	Invitrogen	Cat# 31430 (1:3000 to 1:5000)

ANTIBODIES	SOURCE	IDENTIFIER/DILUTION
List of primary antibodies used in Immunofluorescence and Immunohistochemistry		
SOX9	Abcam	Cat# ab182579 (Rb) (1:300)
DDX3	Abcam	Cat# ab235940 (Rb) (1:300)
CD44v6	Thermo Fisher Scientific	Cat# BMS125 (Ms) (1:300)
PAF1/PD2 rabbit Ab	Abcam	Cat# ab20662 (Rb)(1:250)
PAF1/PD2 mouse mAb	Our lab	N/A (Ms) (1:100)
Nanog	Santa Cruz Biotechnology	Cat# sc-293121 (Ms) (1:150)
PHF5A	Proteintech	Cat# 15554-1-AP (Rb)(1:200)
Non phospho β-CATENIN	Cell Signaling Technology	Cat# CST-19807P (Rb) (1:300)
ALDH1/2	Santa Cruz	Cat# sc-374149 (Ms) (1:300)
Ki67	Cell Signaling Technology	Cat # 9449 (Ms) (1:300)
List of secondary antibodies used in Immunofluorescence		
Alexa Fluor 568 goat anti-mouse IgG	Life Technologies	Cat# A11004 (1:300)
Alexa Fluor 568 goat anti-rabbit IgG	Life Technologies	Cat# A11011 (1:300)
Alexa Fluor 488 goat anti-mouse IgG	Life Technologies	Cat# A11001 (1:300)
Alexa Fluor 488 goat anti-rabbit IgG	Life Technologies	Cat# A11008 (1:300)
List of antibodies used in Immunoprecipitation assay and ChIP assay		
PAF1/PD2	Abcam	Cat# ab20662 (Rb) (3µg/900µg protein)
PHF5A	Proteintech Group	Cat# 15554-1-AP (Rb)
ChromPure IgG	Jackson Immunoresearch	Cat# 011-000-003 (Rb) (3µg/900µg protein)

	Laboratories	
DDX3	Abcam	Cat# ab235940 (Rb) (3µg/900µg protein)
List of antibody used in Flow cytometry		
CD44 APC Human	Miltenyi Biotech	Cat# 130-098-110