Low expression of KMT2C and KMT2D is associated with improved outcome in Pancreatic Ductal Adenocarcinoma

Joshua B.N. Dawkins*[1], Jun Wang*[1], Eleni Maniati[1], James A. Heward[1], Lola Koniali[1], Hemant M. Kocher[1], Sarah A. Martin[1], Claude Chelala[1], Frances R. Balkwill[1], Jude Fitzgibbon□[1], and Richard P. Grose□[1]

*Joint first authorship; □Joint senior authorship

Authors' Affiliations:

[1]Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ UK

Running title:

KMT2C and KMT2D in Pancreatic Ductal Adenocarcinoma

Keywords:

PDAC, Survival, Lysine methyltransferase, Epigenetics, Cell-cycle

Financial support:

Cancer Research UK (C587/A12888, C587/A16354), the European Research Council (ERC322566); Barts and The London Charity (467/1307)

Corresponding Author:

Richard Grose, Centre for Tumour Biology, Barts Cancer Institute, Queen Mary University of London, Charterhouse Square, London, EC1M 6BQ, United Kingdom. Phone: 00-44-207-882-3574; Email: r.p.grose@qmul.ac.uk

Conflict of Interest:

The authors disclose no potential conflicts of interest.

Manuscript notes:

Word count: 244 + 4819; Figures and tables: 6 (+ Supp.); References: 52

GEO data link:

http://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?token=evkzqyigtzmvvob&acc=GSE75327

Abstract

Genes encoding the histone H3 lysine 4 methyltransferases KMT2C and KMT2D are subject to deletion and mutation in pancreatic ductal adenocarcinoma (PDAC), where these lesions identify a group of patients with a more favorable prognosis. In this study, we demonstrate that low *KMT2C* and *KMT2D* expression in biopsies also defines better outcome groups, with median survivals of 15.9 vs 9.2 months (p = 0.029), and 19.9 vs 11.8 months (p = 0.001) respectively. Experiments with eight human pancreatic cell lines showed attenuated cell proliferation when these methyltransferases were depleted, suggesting that this improved outcome may reflect a cell-cycle block with diminished progression from G_0/G_1. RNA-seq analysis of PDAC cell lines following *KMT2C* or *KMT2D* knockdown, identified 31 and 124 differentially expressed genes respectively, with 19 genes in common. Gene set enrichment analysis revealed significant downregulation of genes related to cell-cycle and growth. These data were corroborated independently by examining *KMT2C/D* signatures extracted from the International Cancer Genome Consortium and The Cancer Genome Atlas datasets. Furthermore, these experiments highlighted a potential role for NCAPD3, a condensin II complex subunit, as an outcome predictor in PDAC using existing gene expression series. Kmt2d depletion in KC/KPC cell lines also led to an increased response to the nucleoside analogue 5-fluorouracil, suggesting that lower levels of this methyltransferase may mediate the sensitivity of PDAC to particular treatments. Therefore, it may also be therapeutically beneficial to target these methyltransferases in PDAC, especially in those patients demonstrating higher KTM2C/D expression.

Introduction

Pancreatic ductal adenocarcinomas (PDAC) make up the majority (>90%) of all pancreatic malignancies and are associated with particularly poor overall survival (1). Patients typically present with invasion and metastases at diagnosis, limiting the opportunities for curative surgical resection. The introduction of next-generation sequencing approaches has accelerated our understanding of the recurring coding mutations present in PDAC (2-6). There appears to be a founder population of cells that have accumulated activating mutations in *KRAS* (>90%) (6), alongside loss-of-function mutations in *TP53* (50-75%) (7-10) and *SMAD4* (~55%) (10,11). Additionally, a significant number of other recurring copy number changes and mutations targeting components of the epigenome have been identified, including the histone lysine (K) methyltransferases *KMT2C* (*MLL3*) and *KMT2D* (*MLL2*) (2-6). Intriguingly, these *KMT2C* and *KMT2D* mutations appear to identify a group of patients with better outcome relative to those with wild-type configuration (5), suggesting that depletion of these methyltransferases may either define less aggressive forms of PDAC, or serendipitously improve the efficacy of existing therapies, where the mechanisms underlying this effect are not known.

The KMT2 family of histone lysine methyltransferases consists of KMT2A (MLL1/ALL1), KMT2B (MLL2/MLL4), KMT2C (MLL3/HLAR), KMT2D (MLL2/ALR/MLL4), KMT2E (MLL5), KMT2F (SET1A), KMT2G (SET1B) and KMT2H (ASH1L) (12). These family members, with the exception of KMT2E and KMT2H, act as catalytic subunits within mammalian COMPASS-like complexes to catalyse the addition of methyl groups to a lysine residue on the amino tail of histone H3 (H3K4) (13). H3K4 exists in unmethylated, mono-methylated (H3K4me1), di-methylated (H3K4me2) and tri-methylated (H3K4me3) states, where H3K4me1 is typically associated with enhancers and H3K4me3 with promoters (14). These KMT2 complexes appear to have different substrate specificities to catalyze the formation of H3K4me1 (KMT2C and KMT2D (15,16)), H3K4me1/me2 (KMT2A and KMT2B (17,18)), and H3K4me1/me2/me3 (KMT2F and KMT2G (19)).

Our focus here is restricted to two of these methyltransferases identified as potential key players in PDAC. Loss of KMT2C and KMT2D in cancer is expected to impact upon gene expression, however such changes appear to be cell type dependent, with both negative and positive effects on cell proliferation reported (15,20-28). We

set out to understand how these methyltransferases impact upon PDAC biology, and whether they may present novel opportunities for patient stratification, personalized therapies, or even therapeutic targets.

Methods

Cell lines

Human tumor cell lines PANC-1 and Capan-2, and the immortalized human pancreatic ductal epithelial cell line, HPDE, were cultured in Dulbecco's Modified Eagle's Medium (DMEM; Sigma Aldrich); BxPC-3, SUIT-2, RWP-1 and COLO 357 in RPMI-1640 medium (Sigma Aldrich); and CFPAC-1 cells in Iscove's Modified Dulbecco's Medium (IMDM) with 25mM HEPES (Lonza) and 2mM L-Glutamine (Sigma Aldrich). PANC-1, Capan-2, HPDE, BxPC-3, SUIT-2, RWP-1 and CFPAC-1 were obtained from ATCC. All human cell lines were obtained between 2008 and 2012, and authenticated between 2011 and 2016 using small tandem repeat profiling conducted by LGC standards and ATCC. DT6606, DT6586 and TB32034 cell lines, derived from the *LSL-Kras$^{G12D/+}$;Pdx-1-Cre* (KC) (29) (DT6585 and DT6606) and *LSL-Kras$^{G12D/+}$;LSL-Trp53$^{R172H/+}$;Pdx-1-Cre* (KPC) (30) (TB32043) mouse models of PDAC, were cultured in DMEM (Sigma Aldrich). DT6606, DT6585 and TB320343 were kindly provided by David Tuveson (Cold Spring Harbor Laboratory). For each cell line, medium was supplemented with 10% heat inactivated Fetal Bovine Serum (FBS) (Life Sciences Solutions), 100units/ml penicillin and 100µg/ml streptomycin (Sigma Aldrich). All cell lines were resuscitated from authentic stocks, cultured for less than two months at 37°C and 5% CO_2, and routinely screened for mycoplasma.

RNAi transfection

Cells were plated in 6-well plates with 2.5ml growth medium without penicillin and streptomycin for forward lipid transfection with Silencer® Select siRNAs (Life Sciences Solutions) and Lipofectamine RNAiMAX (Life Sciences Solutions) (Supplementary Table S1). For human cell transfection, a final concentration of 8.3nM per well was achieved by combining 25pmol siRNA with 5µl Lipofectamine in 500µl Opti-MEM I medium (Life Sciences Solutions) for 15 minutes, before adding to wells for 48 hours. The protocol was the same for murine cell transfection, with 150pmol siRNA used instead to achieve a final concentration of 50nM. In all experiments loss of *KMT2D/Kmt2d*, or *KMT2C/Kmt2c*, expression was confirmed by Western blot or real-time PCR.

Western blot

To prepare protein lysates, culture medium was removed, cells were washed twice with Dulbecco's phosphate buffered saline (DPBS), and radioimmunoprecipitation assay (RIPA) buffer (Sigma-Aldrich) (supplemented with protease inhibitor cocktail I (Roche) and 1:100 phosphatase inhibitor cocktail II (Sigma-Aldrich)) was added for 30 minutes on ice. Lysates were harvested with cell scrapers and cellular debris removed by 30 minutes centrifugation at 16000rcf and 4°C. Protein quantification was determined in a bicinchoninic acid (BCA) assay (4% w/v copper (II) sulphate (Sigma-Aldrich) diluted 1:50 in bicinchoninic acid (Sigma-Aldrich)). Samples were run on NuPAGE® Novex® 3-8% Tris-Acetate or 4-12% Bis-Tris gels (Life Sciences Solutions) and transferred to pre-activated PVDF membranes by iBlot dry transfer (Life Sciences Solutions). Membranes were blocked for an hour with 5% w/v bovine serum albumin (Sigma-Aldrich), or 5% w/v nonfat dry milk (Marvel), in Tris-buffered saline with Tween 20 (TBST) and probed with primary antibodies diluted in blocking buffer overnight at 4°C. Membranes were washed with TBST, incubated with rabbit or mouse horseradish peroxidase-conjugated secondary antibodies diluted in blocking buffer, for an hour at room temperature, and washed with TBST (Supplementary Table S2 for primary and secondary antibodies). Staining was visualized by incubation with Amersham enhanced chemiluminescence (ECL), or ECL prime, Western blotting detection reagents (GE Healthcare) and developed using Super Rx X-ray films (Fujifilm) and a Konica Minolta SRX-101A medical film processor. Equivalence of protein loaded for each sample was confirmed by β-actin staining.

Cell proliferation assay

The eight cell lines were seeded in 6-well plates at either $5x10^4$ cells/well (SUIT-2) or $8x10^4$ cells/well (all other cell lines). Following a 48-hour transfection, cells were washed with DPBS and 3ml of fresh penicillin and streptomycin free media were added. Cells were cultured for 72 hours before detachment with trypsin (Sigma Aldrich) and counted using a Vi-Cell XR automated cell viability analyzer (Beckman Coulter).

Flow-cytometric analysis of cell-cycle

Cells seeded in triplicate wells of 6-well plates at either $1.5x10^5$ cells/well (SUIT-2), or $2x10^5$ cells/well (all other cell lines), were transfected for 48 hours. Three wells per

treatment were harvested and analyzed across three time points (48-hours post transfection, 16-hour treatment with 400ng/ml nocodazole, and 24 hours after medium was replaced). Cells were pelleted by five minutes centrifugation at 16000rcf and 4°C and permeabilized with 1ml ice-cold 70% ethanol. Cells were stored at -80°C before washing with ice-cold DPBS and staining with propidium iodide (PI) (50µg/ml PI (Sigma-Aldrich) with 100µg/ml RNase A (Qiagen) in DPBS) for 15 minutes at 37°C. The YG610/20 filter on the Fortessa II flow cytometer was used to examine DNA staining in 30,000 cells.

Real-time PCR analysis

To validate selective knockdown of *KMT2C* or *KMT2D* mRNA by siRNAs, total RNA was isolated from cells, following 48-hour transfection, using the RNeasy Mini Kit with an on-column DNase digestion (Qiagen). Extracted RNA was quantified using a Nanodrop spectrophotometer and reverse transcribed into cDNA using a High-Capacity cDNA Reverse Transcription Kit (Life Sciences Solutions). Duplex real-time PCRs were performed using iTaq™ Universal Probes Supermix (Bio-Rad) and TaqMan® gene expression assays (Life Sciences Solutions) to examine expression of the gene of interest and the 18S ribosomal RNA internal housekeeping standard (Supplementary Table S3). Data were analyzed using a $2^{-\Delta\Delta Ct}$ method (31) to examine relative expression of each gene, where mRNA expression levels were normalized to levels of 18S rRNA for the targeted siRNAs, and then expressed relative to normalized mRNA expression levels of the control siRNA treatment.

RNA sequencing and analysis

Material for RNA sequencing (RNA-seq) was generated from cells seeded at either $1x10^5$ cells/well (SUIT-2) or $1.6x10^5$ cells/well (PANC-1 and COLO 357) in 6-well plates and transfected for 48 hours. Total RNA was isolated (as above), quantified and integrity measured using a RNA 6000 nano assay kit on a 2100 BioAnalyzer (Agilent). A Ribo-Zero Gold kit (Illumina) was used to deplete rRNA from RNA samples and cDNA libraries were prepared using the TruSeq stranded total RNA library preparation kit (Illumina). Libraries prepared for sequencing were validated on the 2100 BioAnalyzer with a DNA 1000 kit (Agilent) and again after randomizing the samples into seven pools with a high sensitivity DNA analysis kit (Agilent). Paired end sequences (reads) of 100 bp in length were generated using seven lanes of a

HiSeq 2000 (Illumina). RNA-seq data have been deposited in Gene Expression Omnibus (GEO) under the accession number GSE75327.

After FASTQ data quality check using FastQC, raw reads were aligned to the reference genome hg19 using Tophat2 (32). An average of 27.2M aligned paired-end reads (range 19.7-36.4M), corresponding to an average of 75.0% (range 59.2-82.2%) concordant pair alignment rate, were reported (Supplementary Table S4). The number of reads uniquely aligned (mapping quality score $q > 10$) to the exonic region of each gene were counted using HTSeq (33), based on the Ensembl annotation (version 74). *KMT2C* and *KMT2D* siRNA datasets were first analysed independently. Only genes that achieved at least one count per million (CPM) mapped reads in at least three samples were included, leading to 15,912 and 15,818 filtered genes in total for the respective *KMT2C* and *KMT2D* siRNA datasets. These genes were classified into 15 RNA species, with protein-coding transcripts representing 81.8% and 82.3%, respectively. Read counts were further normalised using the conditional quantile normalization (cqn) method (34), accounting for gene length and GC content. Differential expression analysis was then performed using the edgeR package (35), employing the generalised linear model (GLM) approach, for each siRNA versus its control pairwise comparison, adjusting for baseline differences between the cell lines, with an additive model design as "model.matrix(~cellline+siRNA_treatment)". For each pairwise comparison, the significantly differentially expressed (DE) genes were selected using a double threshold of false discovery rate (FDR) < 0.05 and an absolute fold change of at least two. Common DE genes were then identified between different siRNA versus control comparisons within the *KMT2C* and *KMT2D* siRNA datasets. On the basis of log2 fold changes of siRNA treated over control for all filtered genes, Gene Set Enrichment Analysis (GSEA) was performed for each comparison using the GSEA tool to identify the canonical pathways gene sets from the Molecular Signatures Database (MSigDB-C2 v5.0) (36). The gene ontology (GO) biological process (BP) enrichment analysis was also performed for DE genes using the PANTHER classification system (37).

Statistical analysis of clinical and expression data

Two large PDAC datasets, the International Cancer Genome Consortium (ICGC) (38) and The Cancer Genome Atlas (TCGA) (3), with both gene expression and clinical follow-up data available, were used to complete survival analysis. Data on 87 patients

from the ICGC dataset were previously compiled and processed (39). Level 3 gene expression data for TCGA dataset were downloaded via TCGA data portal (https://tcga-data.nci.nih.gov/tcga/). Only annotated and confirmed PDAC patients were selected (108 in total). RNA-seq by Expectation-Maximization (RSEM) normalized expression data for 20,501 genes were obtained. For each gene, low and high expression groups were determined using the method described previously (40). Briefly, each percentile of expression between lower and upper quartiles was used in the Cox regression analysis and the best performing threshold of percentile was determined. As an example, the selection of high and low expression groups for *KMT2C* and *KMT2D* in the ICGC and TCGA datasets are shown (Supplemental Fig. S1a).

Survival modeling and Kaplan-Meier (KM) analysis was undertaken using R statistical environment ("survival" package). Overall survival (OS) was defined as time from diagnosis to death, or to the last follow-up date for survivors. Log-rank test was used to calculate the KM P-values. The Cox proportional hazards model was fitted to every gene independently.

Two additional PDAC gene expression profile (GEP) and clinical follow-up datasets, namely "Stratford" and "BCI_Zhang_merged", compiled and processed previously (39), were also included for validation studies.

Cell chemotherapy response assay

The three murine cell lines (DT6606, DT6585 and TB32043) were seeded in 6-well plates at a density of $4x10^4$ cells/well. Following 48-hour transfection, cells were washed with DPBS, detached with trypsin, and re-plated in 96-well plates at $1x10^4$ cells/well (DT6585 and TB32043), or $5x10^3$ cells/well (DT6606). Following adherence, medium was replaced with DMEM containing different concentrations of 5-fluorouracil (5FU) (Accord Healthcare). After 72 hours incubation, WST-1 reagent (Roche) was added to each well and the optical density (OD) measured at 440nm (reference 630nm) after three hours. To generate log-dose response curves, percentage of cell viability was calculated using the maximal OD as 100% viability.

Results

Decreased *KMT2C/D* expression correlates with favorable outcome in PDAC patients

Inactivation of *KMT2C* and *KMT2D* arises through a combination of gene deletion and/or mutation in PDAC (2-6). In order to assess whether expression levels of these methyltransferases are also linked to patient outcomes, we used gene expression profile (GEP) data from the ICGC and TCGA patient series to compare clinical features of patients with tumors expressing different levels of *KMT2C* and *KMT2D*. In the ICGC dataset, we observed that low levels of *KMT2C* and *KMT2D* expression were independently associated with better OS (*KMT2D*, median 19.9 vs. 11.8 months, log-rank p =0.001, Fig. 1a; *KMT2C*, median 15.9 vs 9.2 months, log-rank p = 0.029, Fig. 1b). Combined low-level expression of *KMT2C* and *KMT2D* also correlated with longer survival (median 15.9 vs 9.2 months, log-rank p = 0.044, Supplemental Fig. S1b). A similar trend was also observed in TCGA dataset, for *KMT2C* and combined *KMT2C/D* expression, however these results did not reach statistical significance (Supplemental Fig. S1c). Overall, these observations are consistent with recent genetic data demonstrating that 12 out of 101 patients with *KMT2C* or *KMT2D* loss-of-function mutations have a superior outcome compared to patients with a wild-type configuration (5).

Silencing of *KMT2C/D* expression reduces proliferation

To determine the cellular effects of depleting *KMT2C* and *KMT2D*, three siRNAs were used to silence both methyltransferases in a panel of eight human pancreatic cell lines; three derived from primary tumors (PANC-1, BxPC-3, Capan-2), four from metastatic sites (SUIT-2, RWP-1, CFPAC-1, COLO 357), and one immortalised from the human pancreatic ductal epithelium (HPDE) (Supplemental Fig. S2a,b). *KMT2D* silencing by two siRNAs located to different regions of the transcript (exons 39 and 48) resulted in a reduction of KMT2D protein (Fig. 2a) and, consistent with previous studies in medulloblastoma and colorectal cancer cells (15), reduced cell proliferation significantly in all cell lines tested (Fig. 2b). A third siRNA targeting *KMT2D* produced a less pronounced effect on growth inhibition and was not included in further analyses. *KMT2C* silencing resulted in significant, albeit less pronounced,

reductions in cell proliferation for the three cell lines tested (Supplemental Fig. S2c,d).

To examine whether these changes in cell proliferation were accompanied by cell-cycle anomalies, analysis of *KMT2D* silenced cells was performed using propidium iodide and a G2/M blocking agent, nocodazole. In these experiments, cells treated with control siRNA accumulated within the G2/M phases, contrasting with an observed G0/G1 block in *KMT2D* depleted cells (Fig. 3a). When the cell-cycle profile of PANC-1 cells was evaluated over several days at 24-hour intervals, *KMT2D* silencing led to a decrease in the number and proportion of cells in G0/G1 with a concomitant increase in the apoptotic fraction (Fig. 3b, Supplemental Fig. S2e). Taken together, these experiments suggest that KMT2D contributes towards normal cell-cycle progression, with loss of this methyltransferase leading to cell-cycle arrest and subsequent apoptosis.

KMT2C and *KMT2D* share overlapping gene signatures

To better understand the mechanism(s) responsible for the cellular effects of KMT2C/D, we assessed the transcriptional changes in response to *KMT2C* and *KMT2D* depletion in three PDAC cell lines (PANC-1, SUIT-2 and COLO 357). These cell lines were selected to include a range of background mutations and both primary and metastatic tumors. *KMT2C* and *KMT2D* were targeted individually across the three cell lines using siRNAs, with each targeted siRNA resulting in a measurable and specific reduction of *KMT2C* or *KMT2D* expression, compared to the scrambled control (Fig. 4a,b). Total RNA was isolated 48 hours post-siRNA treatment and transcriptomes were assessed by RNA-seq using 100 bp paired-end deep sequencing (Supplementary Table S4). Resultant RNA-seq profiles confirmed knockdown of both *KMT2C* and *KMT2D* methyltransferases in their corresponding experiments (Supplemental Fig. S3a,b).

We focussed our attention on differentially expressed protein-coding genes shared between the individual siRNAs and all three PDAC cell lines, noting correlations in log2 fold changes for all quantified genes within the *KMT2C* and *KMT2D* siRNA datasets between different siRNAs compared to control (correlation coefficients ranging from 0.39 to 0.60, Supplemental Fig. S3c). Since our data support a greater role for KMT2D than KMT2C, in our analysis we report data in this order. Taking each methyltransferase separately to start, we observed a total of 567 and 759

DE genes for *KMT2D* siRNAs 1 and 2 respectively, with 124 genes common to both siRNAs and the three lines (Fig. 4c,d) (Supplementary Table S5). Fold changes for the DE genes ranged from 0.14 (downregulated, *C2orf54*) to 24.43 (upregulated, *KRT6B*), with 40 genes exhibiting a decrease, and 84 an increase, in gene expression. *KMT2C* silencing resulted in differential expression across 790 genes, with 31 common across all three siRNAs tested (Fig. 4c,d) (Supplementary Table S6). For these 31 common DE genes, the fold changes ranged from 0.18 (downregulated, *AKR1B10*) to 6.51 (upregulated, *ANO3*), with 27 genes demonstrating an increase, and four a decrease, in expression. There was a striking overlap between DE genes in the *KMT2C* and *KMT2D* siRNA datasets, with 19 genes common to both methyltransferases, each demonstrating consistent directional and fold changes between the two methyltransferases (Supplementary Table S6).

Four of these genes were selected for validation based on their previous associations with pancreatic cancer (*c-MET* (2,41), *Calumenin* (42), *Claudin-1* (43,44) and *PTPN14* (45)), with Western blot analysis confirming expression changes of each at the protein level (Fig. 4e, Supplemental Fig. S4a). *ABCB1*, which encodes an ATP-dependent efflux pump and was decreased by *KMT2D* siRNA treatment, was also included in the validation panel, as its expression has been linked with an increase in drug resistance in pancreatic cancer (46). Decreased expression of *ABCB1*, identified by RNA-seq to be specific to *KMT2D* silencing in the two metastatic cell lines, was confirmed by Western blot analysis (Supplemental Fig. S4b). Global levels of H3K4me1/me2/3 remained largely unchanged at 48 hours after KMT2D siRNA transfection in PANC-1 cells, however after 120 hours, KMT2D depletion was accompanied by a global reduction in H3K4me1/3 and, to a lesser extent, H3K4me2 and total H3 (Supplemental Fig. S4c).

Analysis of pathways associated with silenced *KMT2C* or *KMT2D*

In light of the overlapping data between the transcriptomes of KMT2C and KMT2D depleted lines, GSEA was employed to explore the significant differences between targeted and control siRNAs in all curated canonical pathways. In support of our proliferation and cell-cycle studies, six pathways (cell-cycle, cell-cycle mitotic, DNA replication, DNA repair, mitotic M-M/G1 phases and Fanconi pathways) were significantly downregulated (FDR < 0.05) in all siRNA experiments compared to controls (Fig. 5a,b). Other pathways relating to cell-cycle checkpoints and apoptosis

were also affected, although significance was not reached for all five pairwise comparisons. Three pathways of note (telomere maintenance, meiotic recombination and chromosome maintenance) were highly downregulated in the *KMT2D* siRNA datasets, but to a lesser extent in the *KMT2C* siRNA datasets (Fig. 5a,b).

Clinical correlation of expression of targeted genes

Using the ICGC and TCGA human pancreatic cancer datasets, we next assessed the biological process enrichment of genes that strongly correlated (Pearson's correlation $p < 0.001$) with *KMT2C/D* expression levels (Supplementary Table S7). To this end, the PANTHER classification system identified a significant enrichment of the cell-cycle (adjusted (adj.) p = 1.45e-02 for ICGC dataset; adj. p = 1.02e-06 for TCGA), mitosis (adj. p = 1.62e-03 for TCGA), and DNA repair (adj. p = 1.61e-02 for TCGA) pathways for genes that positively correlated with combined *KMT2C/D* expression. In addition, translation was found to be the most overrepresented biological pathway for genes, where expression showed significant negative correlation with *KMT2C/D* (adj. p = 2.21e-04 for ICGC and adj. p = 5.76e-30 for TCGA). The pathways observed to correlate with *KMT2C/D* expression in human pancreatic cancer datasets were therefore consistent with the most significantly downregulated and upregulated pathways following *KMT2C/D* knockdown in our PDAC cell lines (Fig. 5a).

Among the 94 and 257 cell-cycle genes (significantly positively correlated with combined *KMT2C/D* expression) from the ICGC and TCGA datasets, respectively (Supplementary Table S8), we identified three (*NCAPD3*, *CDKL1* and *EIF2AK4*) as being significantly downregulated in at least one *KMT2C/D* siRNA experiment (Supplementary Table S9). Furthermore, patients with low-level expression of these three genes were found to have better OS rates in both the ICGC and TCGA datasets (Fig. 5c and Supplemental Fig. S5a). We focused primarily on NCAPD3, a subunit of the condensin II protein complex involved in chromosome condensation, due to its significant positive correlations with *KMT2C/D* expression in the ICGC and TCGA datasets (r = 0.36, p = 5.56e-04 and r = 0.38, p = 5.19e-05, respectively, Supplementary Fig. S5b), and significant associations with clinical outcome in two additional human pancreatic cancer cohorts (log-rank P = 0.042 for Stratford dataset and P = 2.12e-04 for BCI_Zhang_merged dataset, Supplementary Fig. S5c). The candidacy of *CDKL1* and *EIF2AK4* was not supported in these two additional cohorts (Supplementary Fig. S5c). Interestingly, *NCAPD3* showed

significant downregulation in our *KMT2D* siRNA datasets compared to control (FDR < 0.05), whereas its expression remained almost unchanged in the *KMT2C* siRNA datasets, findings that were confirmed at the protein level (Fig. 5d). These data are consistent with GSEA pathway analysis (Fig. 5b), which highlighted a predominant role for *KMT2D,* but not *KMT2C*, in chromosome maintenance.

Kmt2d depletion increases 5-Fluorouracil sensitivity

The dependence of PDAC cells on KMT2D for survival, and the cell-cycle inhibition renders it an appealing therapeutic target. We postulated that lower expression, and/or mutation, could increase sensitivity to commonly used chemotherapies that target cell-cycle processes, providing one potential mechanism for favourable outcome seen in patients with lower *KMT2D* expression. This hypothesis was tested using murine *Kmt2d* siRNAs in three cell lines derived from the KC (29) (DT6585 and DT6606) and KPC (30) (TB32043) genetically engineered mouse models of pancreatic cancer (Fig. 6a). We were unable to perform these experiments in human cell lines since loss of KMT2D affected their cell proliferation (Fig. 2), something that was not seen in the three murine cell lines (Supplemental Fig. S6a). In both KC and KPC lines we observed an increase in their sensitivity to the nucleoside analogue 5-Fluorouracil (5FU), with exposure to ten-fold less 5FU capable of eliciting the same reduction in cell viability when used in combination with Kmt2d depletion (Fig. 6b). Interestingly, this increased sensitivity to 5FU was specific to Kmt2d loss (Supplemental Fig. S6b,c,d), in accordance with our human data where, relative to KMT2D, *KMT2C* loss had weaker effects on cell proliferation (Supplemental Fig. S2d), and less of an impact on patient survival (ICGC data) (Fig. 1).

Discussion

The identification of mutations in several key methyl and acetyltransferases are now an established feature of many different cancers (47). Mutations in these enzymes are deemed insufficient for cancer initiation alone, with many of these lesions associated with various developmental disorders (48), where mutations in *KMT2D* are responsible for Kabuki syndrome (49). *KMT2C* and *KMT2D* represent a second tier of mutations in PDAC occurring at a lower frequency compared to mutations in *KRAS*, *TP53* and *SMAD4*. The actual frequencies of *KMT2C* and *KMT2D* mutations in PDAC are still debatable (2-6). This is further complicated by different mechanisms of inactivation that include chromosomal deletion and variations in the nature and location of mutations, resulting in missense or truncation, and therefore loss of the functional enzymatic methyltransferase SET domain from the carboxyl-terminal.

In our study, we reasoned that inter-patient fluctuations in expression of these methyltransferases might impart significant changes in outcome. We were intrigued by our initial studies comparing matched GEP and outcome in publically available PDAC series that uncovered a strong favorable signal linked with low expression of these histone modifiers; something that has also been previously reported for *KMT2D* in breast cancer (24). It is reassuring that expression of these methyltransferases is consistent with recent studies that link mutations in these genes, and another family member (*KMT2A*), with improved overall and progression free survival in PDAC (5).

The molecular mechanisms by which these methyltransferases contribute to PDAC development and/or influence patient outcome are likely to be complex, with the H3K4me1-3 chromatin marks potentially altering expression of many target genes. Our silencing experiments, which resulted in near complete loss of these methyltransferases and led to a marked reduction in proliferation effects of *KMT2D* in eight pancreatic cell lines, were consistent with previous studies using gene editing in other tumor settings (15), supporting some overlapping function across tumor types. Indeed our studies go one step further by also implicating KMT2C in cell proliferation, suggesting that these proteins have somewhat complementary roles in PDAC, and may well explain the observation that *KMT2D* and *KMT2C* mutations arise independently of each other (5,50).

We therefore went on to explore the potential downstream effectors of *KMT2C* and *KMT2D* loss in pancreatic cancer, focusing on changes common to both

methyltransferases. Our RNA-seq experiments and subsequent GSEA analysis highlighted changes in DNA replication, repair, and cell-cycle, echoing the results of our initial *in vitro* observations. In order to identify the more robust gene signatures, we focused primarily on changes consistent across the three cell lines, and siRNA experiments. This highlighted several proteins previously implicated in PDAC biology including c-MET, Calumenin, Claudin-1 and PTPN14, all of which were upregulated in our experiments and are known to impact on a range of distinct pathways and processes (41,42,44,51). Indeed, all four were included in the 19-gene signature that was common to *KMT2D* and *KMT2C* in our siRNAs profiles. This gene signature however, did not predict prognosis in any of the four PDAC GEP datasets tested (ICGC, TCGA, Stratford, or BCI_Zhang_merged) based on consensus clustering and OS (data not shown), leading us to speculate that these genes may contribute instead to PDAC development. Moreover, patient clustering based on this 19-gene signature alone did not show a significant enrichment in high or low *KMT2C/D* expression (data not shown). While unexpected, this may reflect the inherent complexity in direct comparisons between primary and cell line data, and may also be a measure of the complete loss of protein induced by siRNA, a feature unlikely to be reflected physiologically, where other compensatory mechanisms might be induced. When we instead focused on genes significantly correlated with *KMT2C/D* expression from the ICGC and TCGA datasets, it was encouraging to find that significant overrepresentation of the cell-cycle signature positively correlated with *KMT2C/D* gene expression. We could refine this cell-cycle compartment by comparing against our RNA-seq expression profile, identifying several common genes with concurrent direction of fold change and magnitude. An examination of the contribution of each gene in turn, and their clinical associations, highlighted a potentially novel role for the NCAPD3 subunit of the condensin II complex in PDAC biology. This chromosome condensation protein demonstrated a 1.3-4.1 fold reduction in expression across the three tested cell lines for *KMT2D*, and encouragingly was a good predictor of outcome in all four GEP series. Having confirmed these findings at the protein level, NCAPD3 is now the focus of ongoing experiments.

Further work remains to determine the mechanisms by which loss or reduced activity of these methyltransferases associates with improved patient outcome. In some pilot experiments we have examined changes in sensitivity to chemotherapy

when Kmt2d, or Kmt2c, were depleted. For these experiments we used cell lines derived from murine models of PDAC whose proliferation, unlike human cell lines, remained unaffected by methyltransferase depletion. Here we noted that Kmt2d silencing increased sensitivity to the antimetabolite 5FU, suggesting that favorable outcome linked with *KMT2D* low expression might be attributable to an improved response to chemotherapy. This effect was not associated with a change in the levels of Abcb1 (data not shown), which has previously been shown to be a mediator of 5FU response (52). This overall effect was specific to *Kmt2d*, as murine cells were not sensitized to 5FU upon *Kmt2c* depletion, perhaps reflecting the weaker impact of low *KMT2C* expression on patient outcome.

In summary, we have identified roles for KMT2D, KMT2C, and a new role for NCAPD3 expression as prognostic predictors in PDAC. The data support the incorporation of combined *KMT2C/D* mutation and gene expression into existing risk stratification models. In addition we report that loss of KMT2D, and to a lesser extent KMT2C, impacts on the cell-cycle and DNA replication pathways, leading to a reduction in cell proliferation. Overall, our studies point to therapeutic benefits of targeting these methyltransferases in PDAC, especially in those patients that demonstrate higher KTM2C/D expression.

Disclosure of potential conflicts of interest

No potential conflicts of interest were disclosed.

Acknowledgements

The authors thank all members of the Fitzgibbon and Balkwill laboratories for advice and guidance of this project. The authors also thank Dave Tuveson and his laboratory for providing the KC and KPC cell lines, and to our other colleagues and collaborators that provided access to the human pancreatic cell lines.

References

1. Hariharan D, Saied A, Kocher HM. Analysis of mortality rates for pancreatic cancer across the world. HPB : the official journal of the International Hepato Pancreato Biliary Association 2008;10:58-62.
2. Waddell N, Pajic M, Patch AM, Chang DK, Kassahn KS, Bailey P, et al. Whole genomes redefine the mutational landscape of pancreatic cancer. Nature 2015;518:495-501.
3. Biankin AV, Waddell N, Kassahn KS, Gingras MC, Muthuswamy LB, Johns AL, et al. Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes. Nature 2012;491:399-405.
4. Jones S, Zhang X, Parsons DW, Lin JC, Leary RJ, Angenendt P, et al. Core signaling pathways in human pancreatic cancers revealed by global genomic analyses. Science 2008;321:1801-6.
5. Sausen M, Phallen J, Adleff V, Jones S, Leary RJ, Barrett MT, et al. Clinical implications of genomic alterations in the tumour and circulation of pancreatic cancer patients. Nature communications 2015;6:7686.
6. Bailey P, Chang DK, Nones K, Johns AL, Patch AM, Gingras MC, et al. Genomic analyses identify molecular subtypes of pancreatic cancer. Nature 2016;531:47-52.
7. Scarpa A, Capelli P, Mukai K, Zamboni G, Oda T, Iacono C, et al. Pancreatic adenocarcinomas frequently show p53 gene mutations. The American journal of pathology 1993;142:1534-43.
8. Redston MS, Caldas C, Seymour AB, Hruban RH, da Costa L, Yeo CJ, et al. p53 mutations in pancreatic carcinoma and evidence of common involvement of homocopolymer tracts in DNA microdeletions. Cancer research 1994;54:3025-33.
9. Barton CM, Staddon SL, Hughes CM, Hall PA, O'Sullivan C, Kloppel G, et al. Abnormalities of the p53 tumour suppressor gene in human pancreatic cancer. British journal of cancer 1991;64:1076-82.
10. Wilentz RE, Iacobuzio-Donahue CA, Argani P, McCarthy DM, Parsons JL, Yeo CJ, et al. Loss of expression of Dpc4 in pancreatic intraepithelial neoplasia: evidence that DPC4 inactivation occurs late in neoplastic progression. Cancer research 2000;60:2002-6.

11. Hahn SA, Schutte M, Hoque AT, Moskaluk CA, da Costa LT, Rozenblum E, et al. DPC4, a candidate tumor suppressor gene at human chromosome 18q21.1. Science 1996;271:350-3.
12. Allis CD, Berger SL, Cote J, Dent S, Jenuwien T, Kouzarides T, et al. New nomenclature for chromatin-modifying enzymes. Cell 2007;131:633-6.
13. Shilatifard A. The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. Annual review of biochemistry 2012;81:65-95.
14. Pekowska A, Benoukraf T, Zacarias-Cabeza J, Belhocine M, Koch F, Holota H, et al. H3K4 tri-methylation provides an epigenetic signature of active enhancers. The EMBO journal 2011;30:4198-210.
15. Guo C, Chen LH, Huang Y, Chang CC, Wang P, Pirozzi CJ, et al. KMT2D maintains neoplastic cell proliferation and global histone H3 lysine 4 monomethylation. Oncotarget 2013;4:2144-53.
16. Lee JE, Wang C, Xu S, Cho YW, Wang L, Feng X, et al. H3K4 mono- and di-methyltransferase MLL4 is required for enhancer activation during cell differentiation. Elife 2013;2:e01503.
17. Wu L, Lee SY, Zhou B, Nguyen UT, Muir TW, Tan S, et al. ASH2L regulates ubiquitylation signaling to MLL: trans-regulation of H3 K4 methylation in higher eukaryotes. Molecular cell 2013;49:1108-20.
18. Patel A, Vought VE, Dharmarajan V, Cosgrove MS. A conserved arginine-containing motif crucial for the assembly and enzymatic activity of the mixed lineage leukemia protein-1 core complex. The Journal of biological chemistry 2008;283:32162-75.
19. Shilatifard A. Chromatin modifications by methylation and ubiquitination: implications in the regulation of gene expression. Annual review of biochemistry 2006;75:243-69.
20. Lee J, Kim DH, Lee S, Yang QH, Lee DK, Lee SK, et al. A tumor suppressive coactivator complex of p53 containing ASC-2 and histone H3-lysine-4 methyltransferase MLL3 or its paralogue MLL4. Proceedings of the National Academy of Sciences of the United States of America 2009;106:8513-8.
21. Kanda H, Nguyen A, Chen L, Okano H, Hariharan IK. The Drosophila ortholog of MLL3 and MLL4, trithorax related, functions as a negative regulator of tissue growth. Molecular and cellular biology 2013;33:1702-10.
22. Santos MA, Faryabi RB, Ergen AV, Day AM, Malhowski A, Canela A, et al. DNA-damage-induced differentiation of leukaemic cells as an anti-cancer barrier. Nature 2014;514:107-11.
23. Chen C, Liu Y, Rappaport AR, Kitzing T, Schultz N, Zhao Z, et al. MLL3 is a haploinsufficient 7q tumor suppressor in acute myeloid leukemia. Cancer cell 2014;25:652-65.
24. Kim JH, Sharma A, Dhar SS, Lee SH, Gu B, Chan CH, et al. UTX and MLL4 coordinately regulate transcriptional programs for cell proliferation and invasiveness in breast cancer cells. Cancer research 2014;74:1705-17.
25. Issaeva I, Zonis Y, Rozovskaia T, Orlovsky K, Croce CM, Nakamura T, et al. Knockdown of ALR (MLL2) reveals ALR target genes and leads to alterations in cell adhesion and growth. Molecular and cellular biology 2007;27:1889-903.

26. Zhu J, Sammons MA, Donahue G, Dou Z, Vedadi M, Getlik M, et al. Gain-of-function p53 mutants co-opt chromatin pathways to drive cancer growth. Nature 2015;525:206-11.
27. Zhang J, Dominguez-Sola D, Hussein S, Lee JE, Holmes AB, Bansal M, et al. Disruption of KMT2D perturbs germinal center B cell development and promotes lymphomagenesis. Nature medicine 2015;21:1190-8.
28. Matkar S, Sharma P, Gao S, Gurung B, Katona BW, Liao J, et al. An Epigenetic Pathway Regulates Sensitivity of Breast Cancer Cells to HER2 Inhibition via FOXO/c-Myc Axis. Cancer cell 2015;28:472-85.
29. Hingorani SR, Petricoin EF, Maitra A, Rajapakse V, King C, Jacobetz MA, et al. Preinvasive and invasive ductal pancreatic cancer and its early detection in the mouse. Cancer cell 2003;4:437-50.
30. Hingorani SR, Wang L, Multani AS, Combs C, Deramaudt TB, Hruban RH, et al. Trp53R172H and KrasG12D cooperate to promote chromosomal instability and widely metastatic pancreatic ductal adenocarcinoma in mice. Cancer cell 2005;7:469-83.
31. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 2001;25:402-8.
32. Kim D, Pertea G, Trapnell C, Pimentel H, Kelley R, Salzberg SL. TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome biology 2013;14:R36.
33. Anders S, Pyl PT, Huber W. HTSeq--a Python framework to work with high-throughput sequencing data. Bioinformatics 2015;31:166-9.
34. Hansen KD, Irizarry RA, Wu Z. Removing technical variability in RNA-seq data using conditional quantile normalization. Biostatistics 2012;13:204-16.
35. Robinson MD, McCarthy DJ, Smyth GK. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139-40.
36. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, et al. Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proceedings of the National Academy of Sciences of the United States of America 2005;102:15545-50.
37. Mi H, Dong Q, Muruganujan A, Gaudet P, Lewis S, Thomas PD. PANTHER version 7: improved phylogenetic trees, orthologs and collaboration with the Gene Ontology Consortium. Nucleic Acids Res 2010;38:D204-10.
38. Zhang J, Baran J, Cros A, Guberman JM, Haider S, Hsu J, et al. International Cancer Genome Consortium Data Portal--a one-stop shop for cancer genomics data. Database (Oxford) 2011;2011:bar026.
39. Haider S, Wang J, Nagano A, Desai A, Arumugam P, Dumartin L, et al. A multi-gene signature predicts outcome in patients with pancreatic ductal adenocarcinoma. Genome Med 2014;6:105.
40. Mihaly Z, Kormos M, Lanczky A, Dank M, Budczies J, Szasz MA, et al. A meta-analysis of gene expression-based biomarkers predicting outcome after tamoxifen treatment in breast cancer. Breast cancer research and treatment 2013;140:219-32.

41. Neuzillet C, Couvelard A, Tijeras-Raballand A, de Mestier L, de Gramont A, Bedossa P, et al. High c-Met expression in stage I-II pancreatic adenocarcinoma: proposal for an immunostaining scoring method and correlation with poor prognosis. Histopathology 2015;67:664-76.
42. Wang Q, Shen B, Chen L, Zheng P, Feng H, Hao Q, et al. Extracellular calumenin suppresses ERK1/2 signaling and cell migration by protecting fibulin-1 from MMP-13-mediated proteolysis. Oncogene 2015;34:1006-18.
43. Kondo J, Sato F, Kusumi T, Liu Y, Motonari O, Sato T, et al. Claudin-1 expression is induced by tumor necrosis factor-alpha in human pancreatic cancer cells. International journal of molecular medicine 2008;22:645-9.
44. Borka K, Kaliszky P, Szabo E, Lotz G, Kupcsulik P, Schaff Z, et al. Claudin expression in pancreatic endocrine tumors as compared with ductal adenocarcinomas. Virchows Archiv : an international journal of pathology 2007;450:549-57.
45. Niedergethmann M, Alves F, Neff JK, Heidrich B, Aramin N, Li L, et al. Gene expression profiling of liver metastases and tumour invasion in pancreatic cancer using an orthotopic SCID mouse model. British journal of cancer 2007;97:1432-40.
46. Pang L, Word B, Xu J, Wang H, Hammons G, Huang SM, et al. ATP-Binding Cassette Genes Genotype and Expression: A Potential Association with Pancreatic Cancer Development and Chemoresistance? Gastroenterology research and practice 2014;2014:414931.
47. Plass C, Pfister SM, Lindroth AM, Bogatyrova O, Claus R, Lichter P. Mutations in regulators of the epigenome and their connections to global chromatin patterns in cancer. Nature reviews Genetics 2013;14:765-80.
48. Fahrner JA, Bjornsson HT. Mendelian disorders of the epigenetic machinery: tipping the balance of chromatin states. Annu Rev Genomics Hum Genet 2014;15:269-93.
49. Ng SB, Bigham AW, Buckingham KJ, Hannibal MC, McMillin MJ, Gildersleeve HI, et al. Exome sequencing identifies MLL2 mutations as a cause of Kabuki syndrome. Nature genetics 2010;42:790-3.
50. Witkiewicz AK, McMillan EA, Balaji U, Baek G, Lin WC, Mansour J, et al. Whole-exome sequencing of pancreatic cancer defines genetic diversity and therapeutic targets. Nature communications 2015;6:6744.
51. Huang JM, Nagatomo I, Suzuki E, Mizuno T, Kumagai T, Berezov A, et al. YAP modifies cancer cell sensitivity to EGFR and survivin inhibitors and is negatively regulated by the non-receptor type protein tyrosine phosphatase 14. Oncogene 2013;32:2220-9.
52. Lu F, Hou YQ, Song Y, Yuan ZJ. TFPI-2 downregulates multidrug resistance protein in 5-FU-resistant human hepatocellular carcinoma BEL-7402/5-FU cells. Anat Rec (Hoboken) 2013;296:56-63.

Figure legends (Main body)

Figure 1

KM survival analysis to assess prognostic value of *KMT2D* (**A**) and *KMT2C* (**B**) expression in the ICGC dataset. Numbers on the x-axis are in the unit of years. Lower expression of *KMT2C* and *KMT2D* in PDAC tumors (black) correlates with improved patient survival compared to higher expression (gray).

Figure 2

Depletion of KMT2D reduces cell proliferation. **A,** Western blot analysis confirmed reduced expression of KMT2D at 72 hours post-transfection, with *KMT2D* targeted siRNA duplexes, in eight cell lines. A third siRNA was cropped from the images. **B,** Depletion of KMT2D significantly inhibits cell proliferation (One-way ANOVA with Dunnett's post-hoc analysis, ** $p < 0.01$, *** $p < 0.001$), where proliferation of the eight cell lines was examined by performing a cell count at 72 hours post *KMT2D* siRNA transfection. Data shown are mean values for three replicate wells performed on the same day ± SD for control siRNA, *KMT2D* siRNA1, and *KMT2D* siRNA2.

Figure 3

Loss of KMT2D expression blocks cells in G0/G1 before cells undergo apoptosis. **A,** Cell-cycle analysis of the SUIT-2 cell line, which is representative of the eight cell lines tested. Profiles and graphs show that cells with reduced KMT2D do not accumulate in G2/M after 16 hours of 400ng/ml nocodazole treatment, due to their retention in the G0/G1 fraction. (U: Untreated, N: Nocodazole, R: Recovery). **B,** Graphs depicting the cell-cycle profiles of PANC-1 cells over 168 hours at 24-hour intervals showing that cells blocked in G0/G1, due to KMT2D loss, begin to undergo apoptosis.

Figure 4

RNA-seq identifies changes in gene expression following loss of *KMT2D* or *KMT2C*. **A-B**, *KMT2D* and *KMT2C* targeted siRNAs reduce expression of *KMT2D* mRNA and protein, and *KMT2C* mRNA respectively. **A,** Western blot analysis showing the two *KMT2D* siRNAs reducing KMT2D levels in three cell lines (PANC-1, SUIT-2 and COLO 357), whereas for the three *KMT2C* and control siRNAs it remains largely unaffected. **B,** RT-qPCR analysis showing changes in *KMT2D* and *KMT2C* mRNA expression in the three cell lines (PANC-1 (black circles), SUIT-2 (red triangles) and

COLO 357 (blue squares)) following transfection with the two *KMT2D* siRNAs, the three *KMT2C* siRNAs, or the scramble negative control siRNA. Data shown are normalized mean values from technical triplicates. **C-D,** Bioinformatic analysis of RNA-seq data identifies DE genes common across the three cell lines for the two *KMT2D* siRNAs, and three *KMT2C* siRNAs, compared to the negative control siRNA. **C,** Venn diagram depicting the commonality for the DE genes of each siRNA across the three cell lines, where 19 were found to be common across all siRNAs and cell lines. **D,** Heat maps for KMT2D and KMT2C datasets showing their 124 and 31 commonly DE genes, respectively. **E,** Western blot analysis validating expression changes of four DE genes at the protein level for the two *KMT2D* siRNAs and three *KMT2C* siRNAs. A third siRNA was cropped from images **A** and **E**.

Figure 5

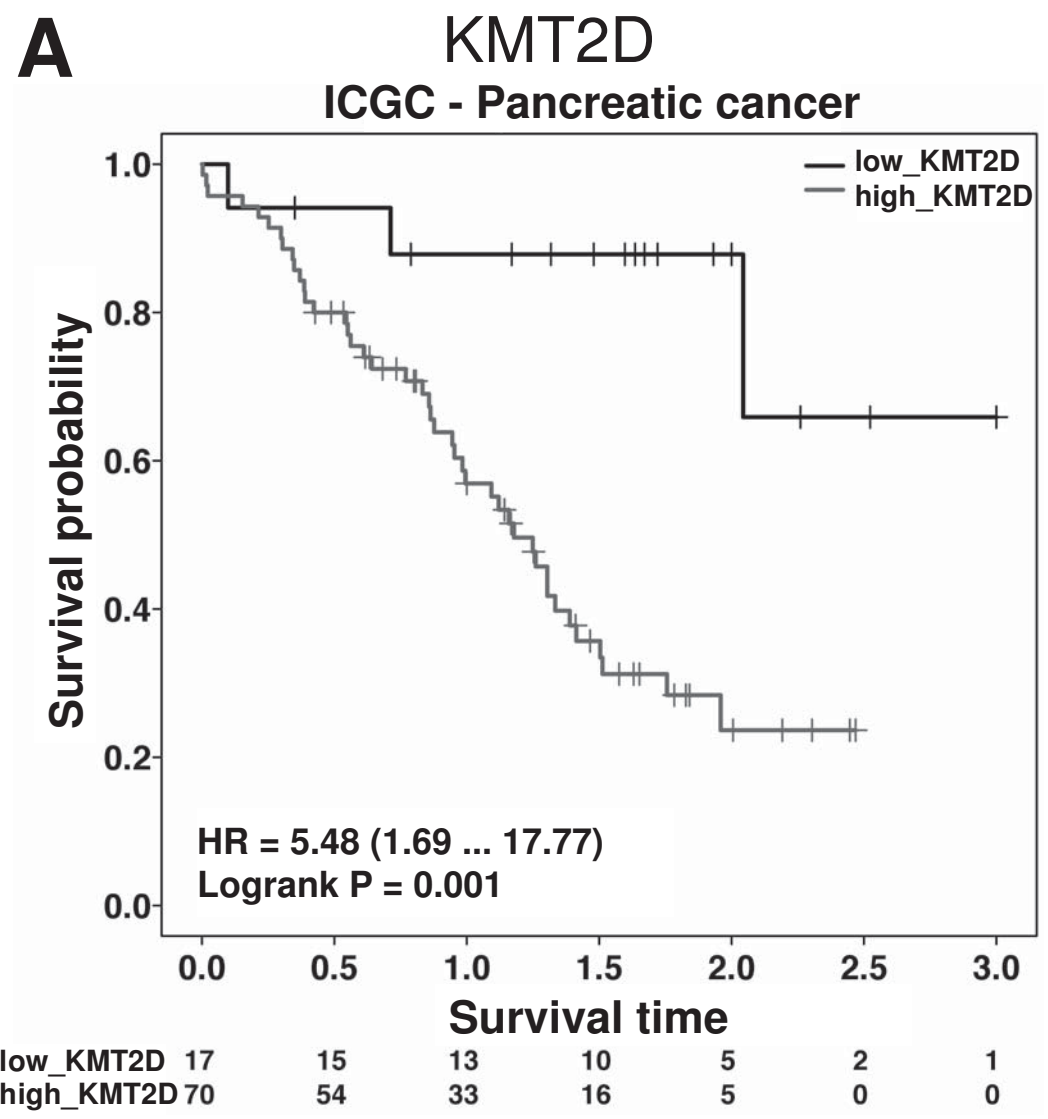
GSEA identifies significant enrichment of genes relating to the cell-cycle, cell growth and DNA repair pathways upon loss of KMT2D, potentially contributing to an improved patient outcome. **A,** Heat map showing changes in the normalized enrichment score (NES) for the upmost enriched pathways following loss of KMT2D or *KMT2C*. A negative NES indicated the down-regulation in siRNA treated samples in relation to control. Pathways significantly downregulated by all examined siRNAs are noted in red. **B,** Selected enrichment score plots for genes in the REACTOME cell-cycle and telomere maintenance pathways. **C,** KM survival analysis of the ICGC and TCGA datasets to show significant negative correlations of patient survival with high (red) and low (black) expression of NCPAD3. Numbers on the x-axis represent years. **D,** Expression analysis data showing significantly reduced *NCAPD3* expression, both by RNA-seq (upper panel) and Western blotting (lower panel), for the two *KMT2D* siRNAs. This reduction was minimal for the three *KMT2C* siRNAs.

Figure 6

Depletion of Kmt2d increases sensitivity of murine pancreatic cancer cells to 5FU. **A,** Western blot analysis showing the reduction in Kmt2d expression by two *Kmt2d* targeted siRNAs in the three murine KC (DT6606 and DT6585) and KPC (TB32043) cell lines. **B,** Reduced Kmt2d expression results in a leftward shift in 5FU dose-cell viability response curves, showing that Kmt2d depletion renders each of the cell lines more sensitive to 5FU treatment. Cell viability was examined using WST-1 after 72 hours of 5FU treatment. Data shown are mean OD values from technical triplicate

wells normalized to maximal OD for each of two biological replicates for control siRNA (Black squares), *Kmt2d* siRNA1 (Light gray circles), and *Kmt2d* siRNA2 (Dark gray triangles).

Figure 1


A
KMT2D
ICGC - Pancreatic cancer
low_KMT2D
high_KMT2D
Survival probability
1.0
0.8
0.6
0.4
0.2
0.0
HR = 5.48 (1.69 ... 17.77)
Logrank P = 0.001
0.0 0.5 1.0 1.5 2.0 2.5 3.0
Survival time
low_KMT2D 17 15 13 10 5 2 1
high_KMT2D 70 54 33 16 5 0 0

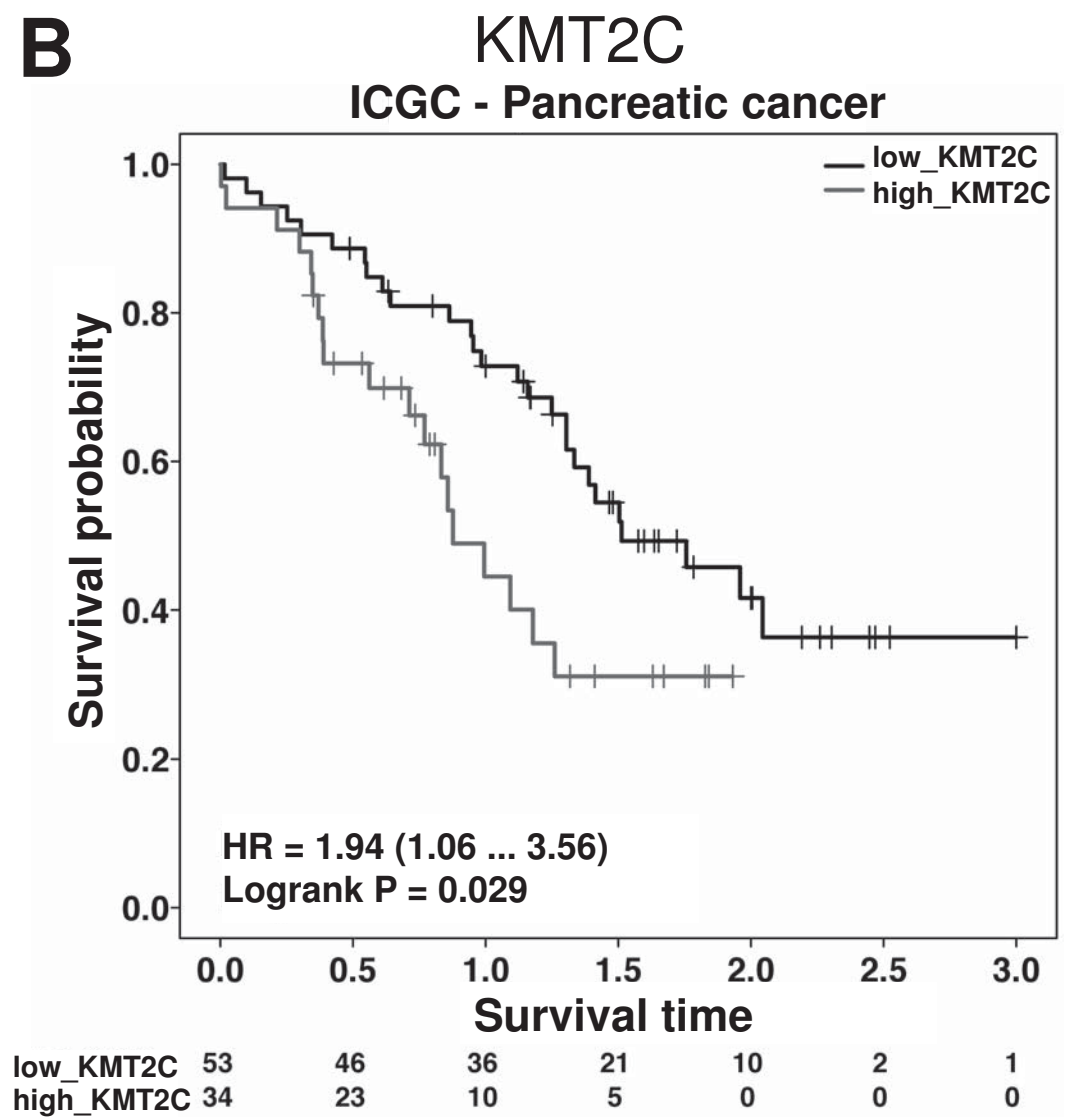
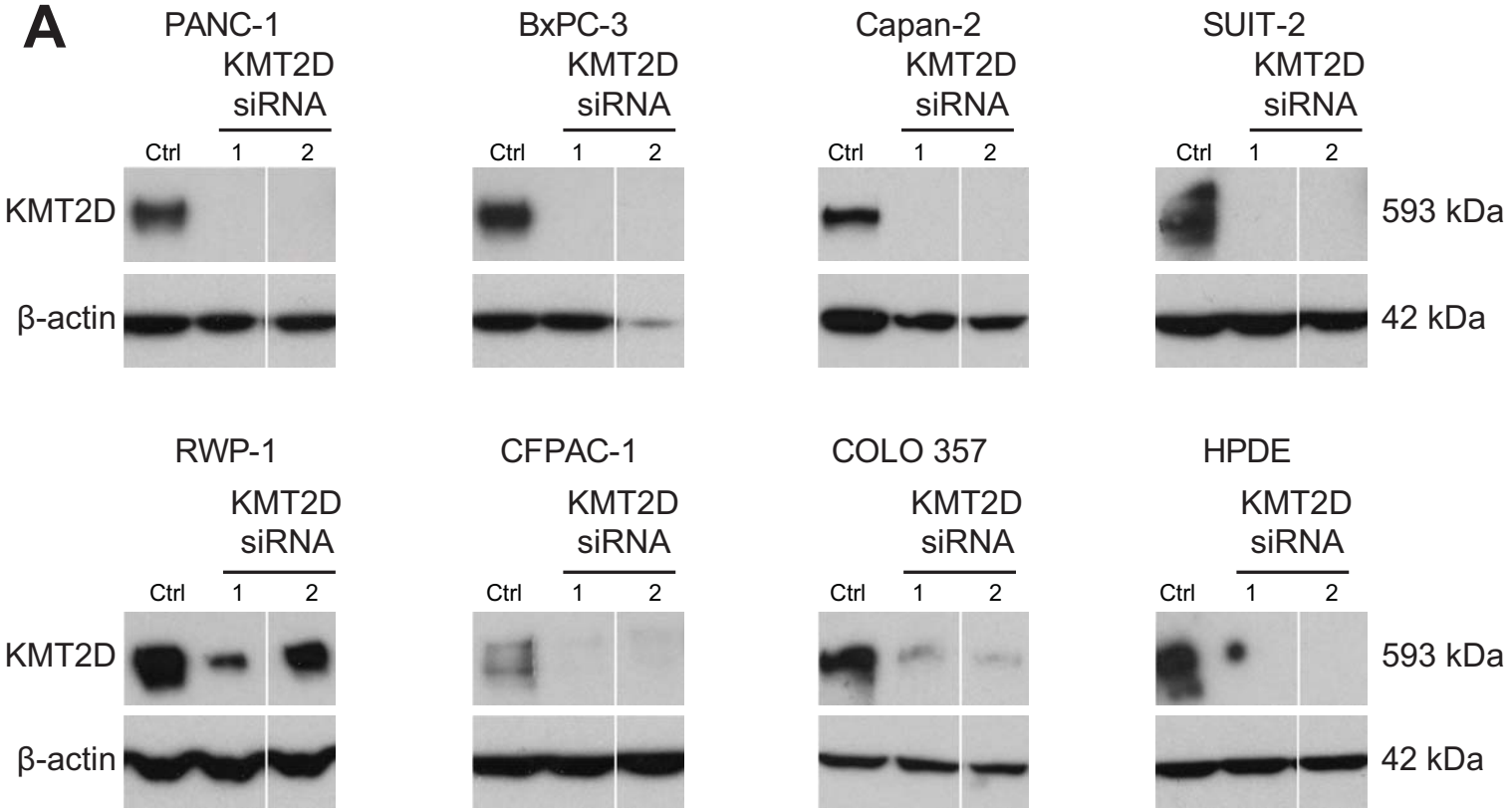

B
KMT2C
ICGC - Pancreatic cancer
low_KMT2C
high_KMT2C
Survival probability
1.0
0.8
0.6
0.4
0.2
0.0
HR = 1.94 (1.06 ... 3.56)
Logrank P = 0.029
0.0 0.5 1.0 1.5 2.0 2.5 3.0
Survival time
low_KMT2C 53 46 36 21 10 2 1
high_KMT2C 34 23 10 5 0 0 0

Figure 2


A
PANC-1
BxPC-3
Capan-2
SUIT-2
KMT2D siRNA
Ctrl
1
2
KMT2D
β-actin
593 kDa
42 kDa
RWP-1
CFPAC-1
COLO 357
HPDE

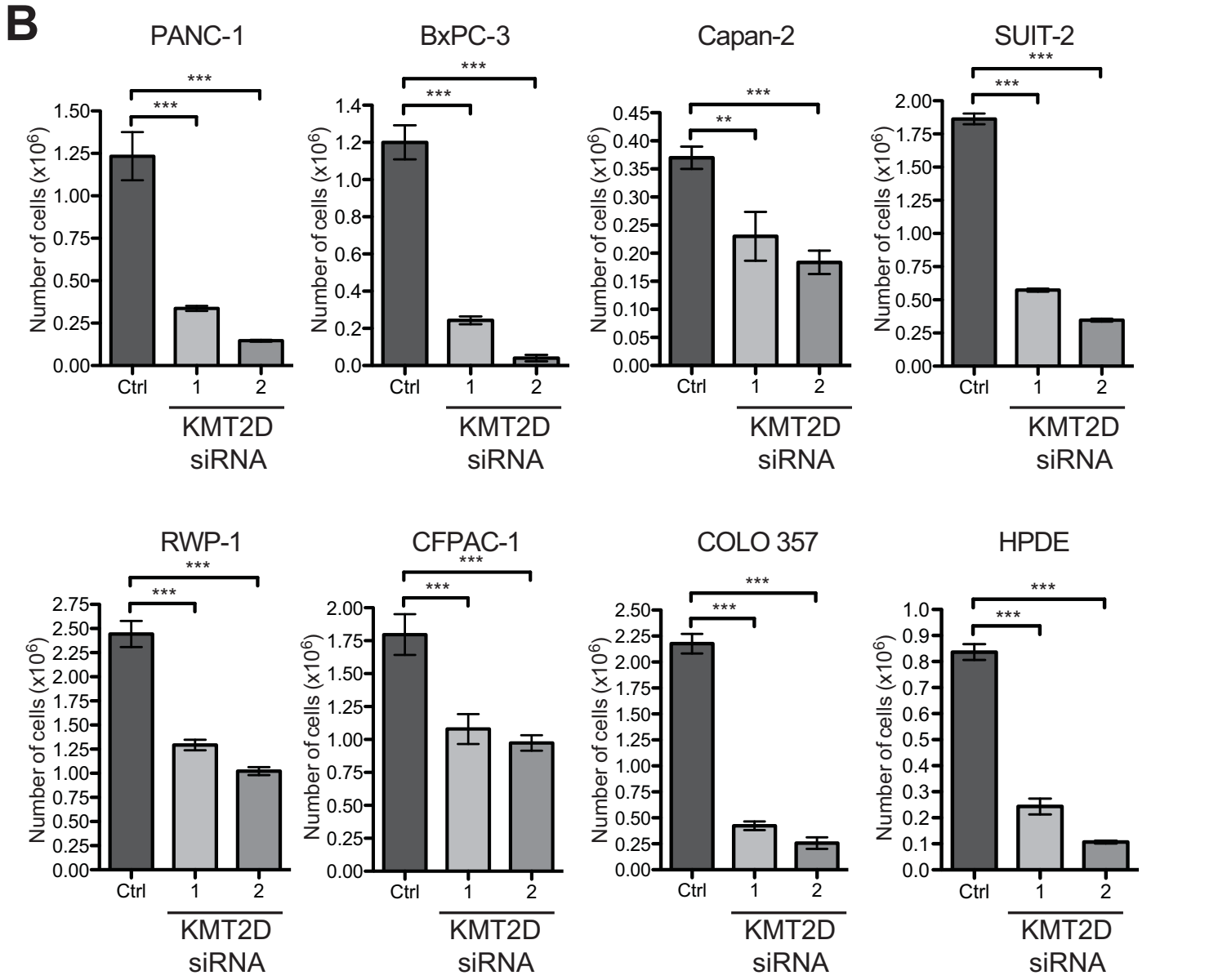

B
PANC-1
BxPC-3
Capan-2
SUIT-2
RWP-1
CFPAC-1
COLO 357
HPDE
Number of cells (x10^6)
***
**
Ctrl
1
2
KMT2D siRNA

Figure 3

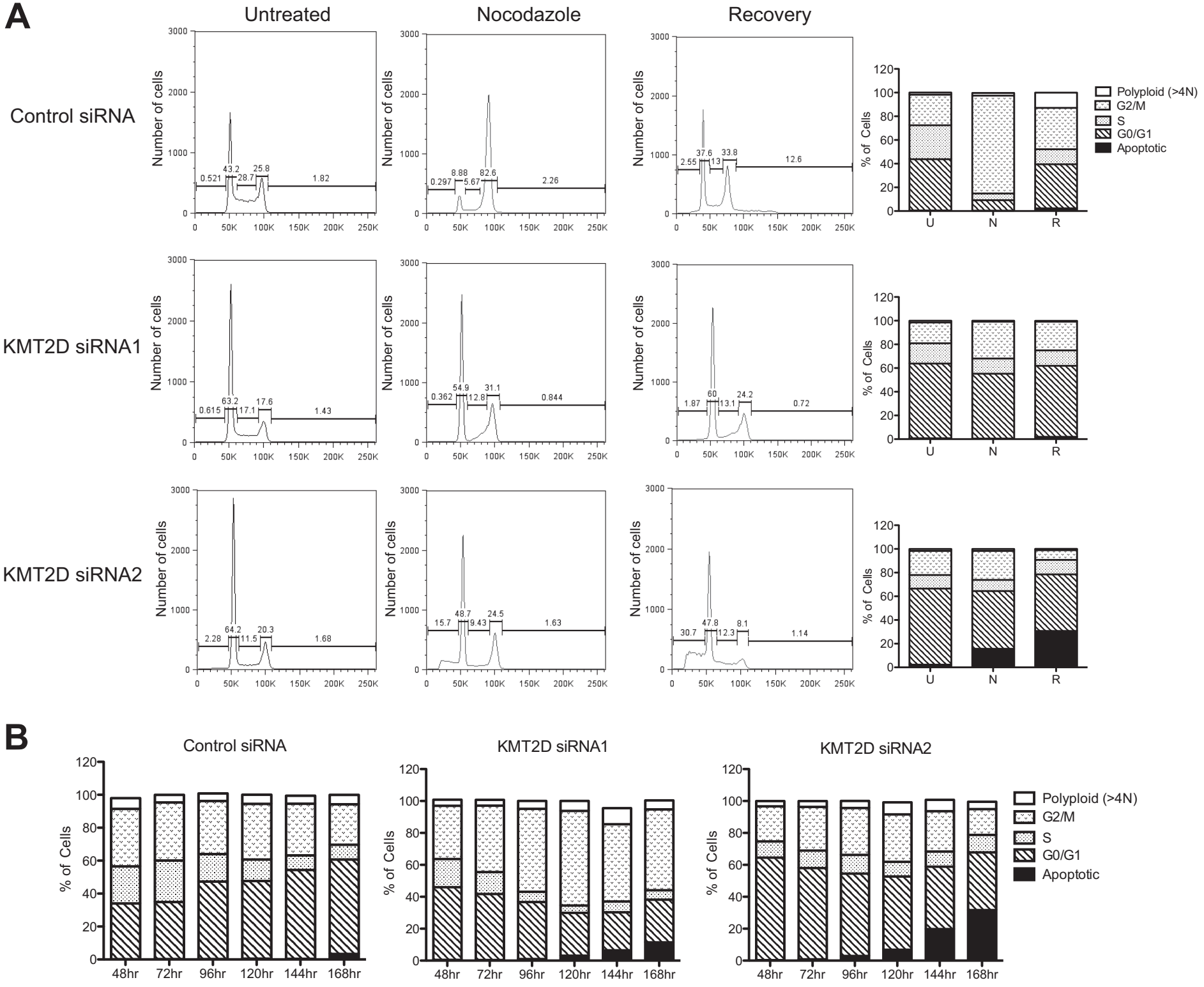

A
Untreated
Nocodazole
Recovery
Control siRNA
KMT2D siRNA1
KMT2D siRNA2
Number of cells
% of Cells
0.521
43.2
28.7
25.8
1.82
0.297
8.88
5.67
82.6
2.26
2.55
37.6
13
33.8
12.6
0.615
63.2
17.1
17.6
1.43
0.362
54.9
12.8
31.1
0.844
1.87
60
13.1
24.2
0.72
2.28
64.2
11.5
20.3
1.68
15.7
48.7
9.43
24.5
1.63
30.7
47.8
12.3
8.1
1.14
U
N
R
Polyploid (>4N)
G2/M
S
G0/G1
Apoptotic
B
Control siRNA
KMT2D siRNA1
KMT2D siRNA2
% of Cells
48hr
72hr
96hr
120hr
144hr
168hr
Polyploid (>4N)
G2/M
S
G0/G1
Apoptotic

Figure 4

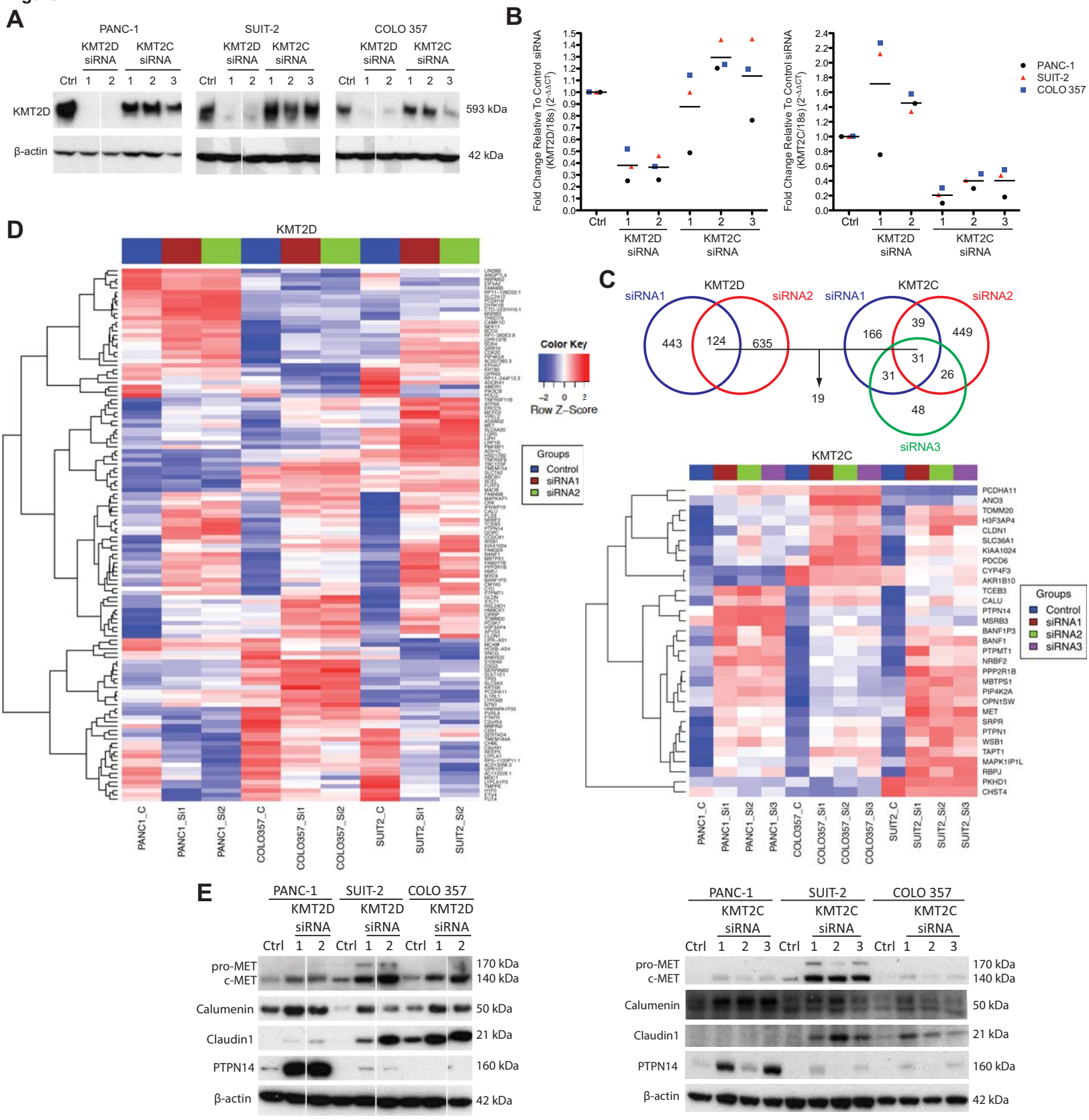

A
PANC-1
SUIT-2
COLO 357
KMT2D siRNA
KMT2C siRNA
Ctrl 1 2 1 2 3
KMT2D
β-actin
593 kDa
42 kDa
B
Fold Change Relative To Control siRNA (KMT2D/18s) (2^-ΔΔCT)
Fold Change Relative To Control siRNA (KMT2C/18s) (2^-ΔΔCT)
PANC-1
SUIT-2
COLO 357
C
siRNA1
siRNA2
siRNA3
KMT2D
KMT2C
443
124
635
166
39
449
31
31
26
19
48
D
KMT2D
Color Key
Row Z-Score
Groups
Control
siRNA1
siRNA2
siRNA3
PANC1_C
PANC1_Si1
PANC1_Si2
PANC1_Si3
COLO357_C
COLO357_Si1
COLO357_Si2
COLO357_Si3
SUIT2_C
SUIT2_Si1
SUIT2_Si2
SUIT2_Si3
E
PANC-1
SUIT-2
COLO 357
KMT2D siRNA
KMT2C siRNA
Ctrl 1 2 3
pro-MET
c-MET
Calumenin
Claudin1
PTPN14
β-actin
170 kDa
140 kDa
50 kDa
21 kDa
160 kDa
42 kDa

Figure 5

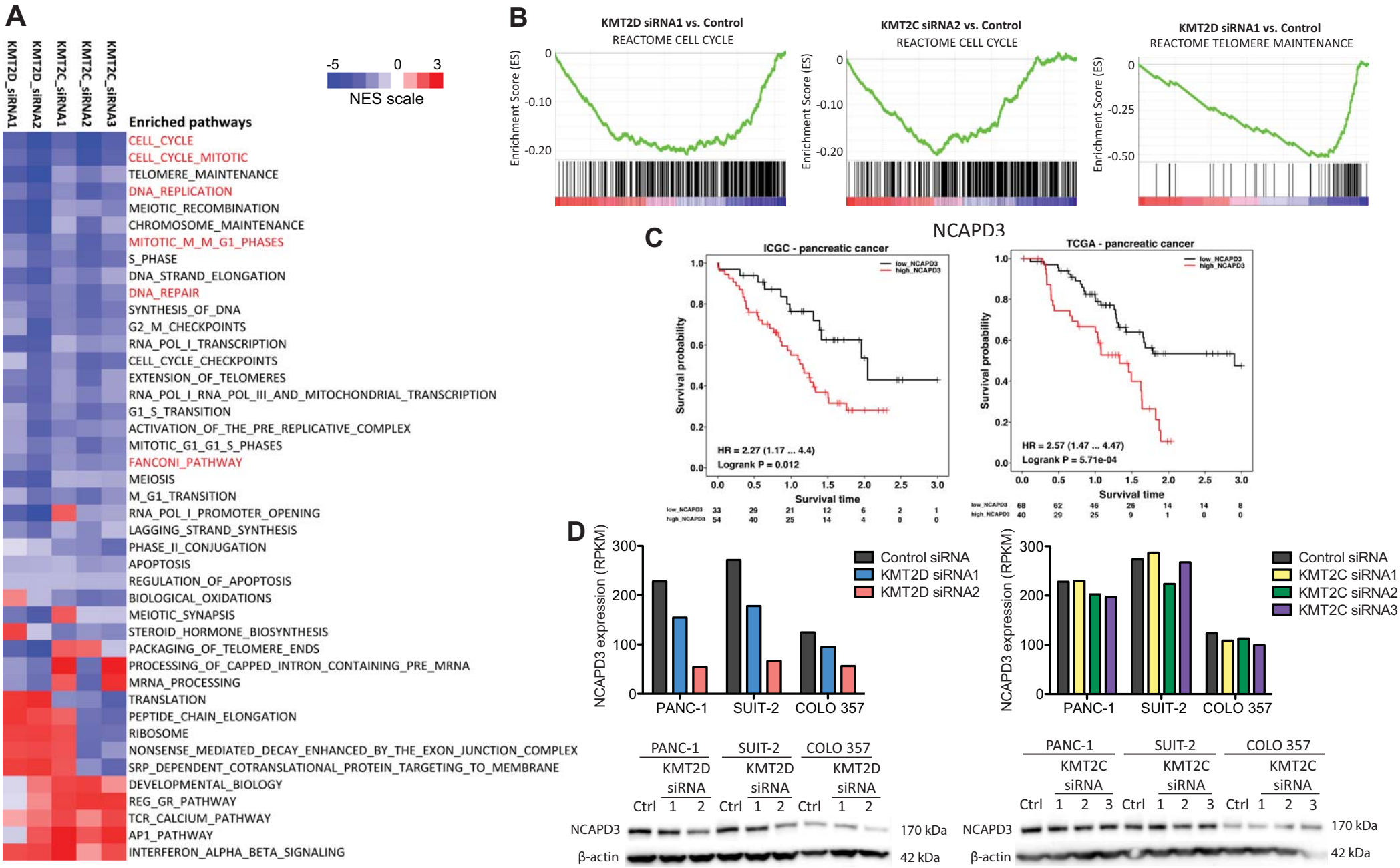

A
KMT2D_siRNA1
KMT2D_siRNA2
KMT2C_siRNA1
KMT2C_siRNA2
KMT2C_siRNA3
-5
0
3
NES scale
Enriched pathways
CELL_CYCLE
CELL_CYCLE_MITOTIC
TELOMERE_MAINTENANCE
DNA_REPLICATION
MEIOTIC_RECOMBINATION
CHROMOSOME_MAINTENANCE
MITOTIC_M_M_G1_PHASES
S_PHASE
DNA_STRAND_ELONGATION
DNA_REPAIR
SYNTHESIS_OF_DNA
G2_M_CHECKPOINTS
RNA_POL_I_TRANSCRIPTION
CELL_CYCLE_CHECKPOINTS
EXTENSION_OF_TELOMERES
RNA_POL_I_RNA_POL_III_AND_MITOCHONDRIAL_TRANSCRIPTION
G1_S_TRANSITION
ACTIVATION_OF_THE_PRE_REPLICATIVE_COMPLEX
MITOTIC_G1_G1_S_PHASES
FANCONI_PATHWAY
MEIOSIS
M_G1_TRANSITION
RNA_POL_I_PROMOTER_OPENING
LAGGING_STRAND_SYNTHESIS
PHASE_II_CONJUGATION
APOPTOSIS
REGULATION_OF_APOPTOSIS
BIOLOGICAL_OXIDATIONS
MEIOTIC_SYNAPSIS
STEROID_HORMONE_BIOSYNTHESIS
PACKAGING_OF_TELOMERE_ENDS
PROCESSING_OF_CAPPED_INTRON_CONTAINING_PRE_MRNA
MRNA_PROCESSING
TRANSLATION
PEPTIDE_CHAIN_ELONGATION
RIBOSOME
NONSENSE_MEDIATED_DECAY_ENHANCED_BY_THE_EXON_JUNCTION_COMPLEX
SRP_DEPENDENT_COTRANSLATIONAL_PROTEIN_TARGETING_TO_MEMBRANE
DEVELOPMENTAL_BIOLOGY
REG_GR_PATHWAY
TCR_CALCIUM_PATHWAY
AP1_PATHWAY
INTERFERON_ALPHA_BETA_SIGNALING
B
KMT2D siRNA1 vs. Control
REACTOME CELL CYCLE
KMT2C siRNA2 vs. Control
REACTOME CELL CYCLE
KMT2D siRNA1 vs. Control
REACTOME TELOMERE MAINTENANCE
Enrichment Score (ES)
C
NCAPD3
ICGC - pancreatic cancer
TCGA - pancreatic cancer
low_NCAPD3
high_NCAPD3
Survival probability
Survival time
HR = 2.27 (1.17 ... 4.4)
Logrank P = 0.012
HR = 2.57 (1.47 ... 4.47)
Logrank P = 5.71e-04
D
NCAPD3 expression (RPKM)
Control siRNA
KMT2D siRNA1
KMT2D siRNA2
KMT2C siRNA1
KMT2C siRNA2
KMT2C siRNA3
PANC-1
SUIT-2
COLO 357
KMT2D siRNA
KMT2C siRNA
Ctrl
NCAPD3
β-actin
170 kDa
42 kDa

Figure 6

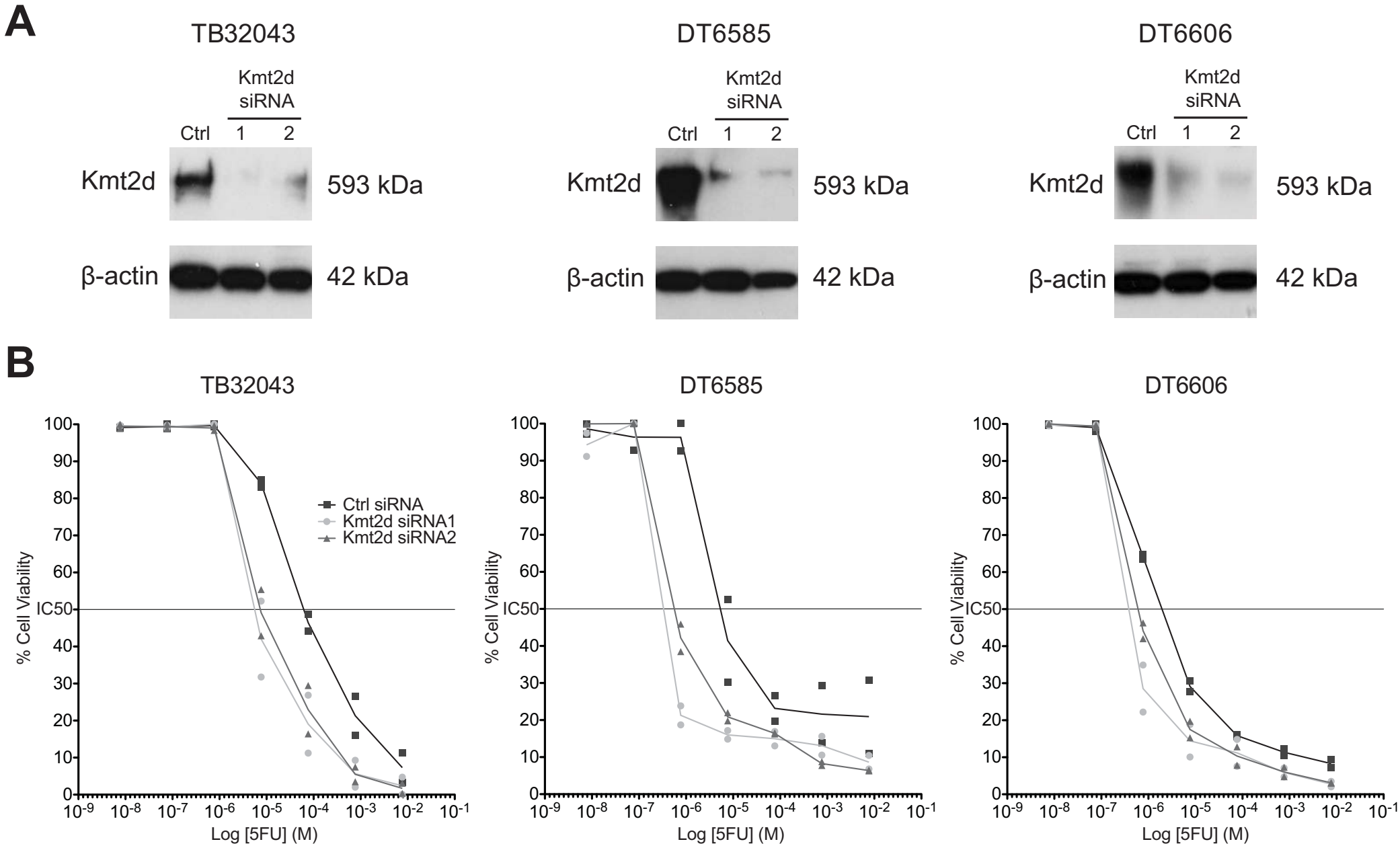

A
TB32043
DT6585
DT6606
Kmt2d siRNA
Ctrl
1
2
Kmt2d
593 kDa
β-actin
42 kDa
B
TB32043
DT6585
DT6606
Ctrl siRNA
Kmt2d siRNA1
Kmt2d siRNA2
% Cell Viability
IC50
Log [5FU] (M)

Figure legends (Supplemental)

Supplementary Figure S1

Analysis of PDAC tumors from the ICGA and TCGA datasets suggests that lower *KMT2C* and *KMT2D* expression may correlate with improved patient survival. **A**, Percentile cut-off plots to determine the best threshold percentile for determining high and low expression groups. **B**, KM survival analysis to assess the prognostic value of combined *KMT2C/D* expression in the ICGC dataset. **C**, KM survival analysis to assess the prognostic value of *KMT2D*, *KMT2C,* and their combined expression in TCGA dataset. Numbers on the x-axis represent years.

Supplementary Figure S2

Expression of KMT2D and *KMT2C*, where *KMT2C* depletion by targeted siRNAs had a significant impact upon cell proliferation and cell cycle. **A**, KMT2D expression was detected by Western Blot across eight human cell lines derived from primary tumors (PANC-1, BxPC-3, Capan-2), metastatic tumors (SUIT-2, CFPAC-1, RWP-1, COLO 357) and an immortalized pancreatic ductal epithelium (HPDE). **B**, Fold changes in *KMT2D* (blue circles) and *KMT2C* (red squares) mRNA expression relative to 18S rRNA were detected by RT-qPCR across the eight cell lines. Data shown are mean values from technical triplicates. **C**, RT-qPCR analysis confirming reduced expression of *KMT2C* mRNA relative to 18S rRNA for the three *KMT2C* siRNAs at 72 hours post-transfection across three cell lines (PANC-1 (black circles), SUIT-2 (red triangles) and COLO 357 (blue squares)). Data shown are normalized mean values from technical triplicates. **D,** *KMT2C* mRNA depletion by targeted siRNAs negatively impacts upon cell proliferation, where the number of cells at 72 hours post-transfection with the three *KMT2C* siRNAs was significantly reduced (One-way ANOVA with Dunnett's post-hoc analysis, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Data shown are mean values for three replicate wells performed on the same day ± SD for control siRNA (gray), *KMT2C* siRNA1 (blue), *KMT2C* siRNA2 (red) and *KMT2C* siRNA3 (green). **E**, Cell-cycle profile analysis from the PANC-1 cells over 168 hours at 24-hour intervals following KMT2D knockdown showing that the cells retained in the G0/G1 fraction begin to undergo apoptosis.

Supplementary Figure S3

Confirmation of *KMT2C* and *KMT2D* knockdown by siRNA at the exon level in RNA-seq data. Expression, normalized counts and exon usage data for *KMT2D* (**A**) and *KMT2C* (**B**) were derived from DEXseq R Bioconductor package following treatment with siRNAs that confirms reduced targeted gene expression.

Supplementary Figure S4

Correlations in RNA-seq data between the multiple siRNAs for each gene. All sufficiently quantified genes were used for the overall correlation between different siRNAs.

Supplementary Figure S5

Changes in RPKM expression for selected genes identified by RNA-seq, and global levels of H3K4 methylation, following depletion of KMT2D or *KMT2C* by siRNA. **A,** RNA-seq RPKM data for four of the 19 commonly DE genes, each increased following *KMT2D* or *KMT2C* siRNA treatment. **B,** RNA-seq RPKM data and Western blot analysis for the decrease in ABCB1 expression across the cell lines following treatment with *KMT2D* siRNAs. A third siRNA was cropped from the images. **C,** Western Blot analysis for H3K4 methylation at 48 and 120 hours following transfection with *KMT2D, KMT2C*, or control siRNAs in PANC-1 cells.

Supplementary Figure S6

Correlations of gene expression and patient survival. **A,** KM survival analysis from the ICGC (upper panel) and TCGA (lower panel) datasets to show significant negative correlations of patient survival with high and low expression of *CDKL1* and *EIF2AK4*. **B,** Significant positive correlations between *NCAPD3* and *KMT2C/D* combined expression levels in both ICGC and TCGA datasets. The Pearson's correlation coefficient and associated p-values are shown for each set. **C,** KM survival analysis to show the gene expression with overall survival in two additional PDAC datasets, Stratford and BCI_Zhang_merged, for *NCAPD3*, *CDKL1* and *EIF2AK4*. Note that *NCAPD3* showed significant negative correlation between expression and overall survival in both sets.

Supplementary Figure S7

Depletion of Kmt2d and *Kmt2c* does not affect the viability of murine pancreatic cancer cells, and *Kmt2c* depletion does not affect sensitivity to 5FU. **A,** Cell viability

was examined by WST-1 at 72 hours post-transfection. Data shown are from mean OD values for 15 replicate wells performed across several days ± SD for untransfected (black), control siRNA (gray), *Kmt2d* siRNA1 (blue), and *Kmt2d* siRNA2 (red). **B,** RT-qPCR analysis showing that the two *Kmt2c* siRNAs reduce *Kmt2c* mRNA expression in three murine cell lines derived from KC (DT6585 (red triangles) and DT6606 (black circles)) and KPC (TB32043 (blue squares)) models of PDAC. **C,** *Kmt2c* siRNAs alone do not have an impact upon cell viability compared to cells not transfected. Data shown are from mean OD values for 15 replicate wells performed across several days ± SD for untransfected (black), control siRNA (gray), *Kmt2c* siRNA1 (blue), and *Kmt2c* siRNA2 (red). **D,** Reduced *Kmt2c* expression by two siRNAs has no impact on the response of the cells to treatment with 5FU. Viability of the three transfected cell lines was examined by WST-1 after 72 hours of treatment with 5FU. Data shown are mean OD values from technical triplicate wells normalized to maximal OD for each of biological replicate (N = 2 for TB32043 and DT6606; N = 1 for DT6585) for control siRNA (gray), *Kmt2c* siRNA1 (blue), and *Kmt2c* siRNA2 (red).

Tables in Supplemental

Supplementary Table S1

Details of the siRNAs used for cell transfections.

Supplementary Table S2

Details of the antibodies and conditions used for Western blots.

Supplementary Table S3

Details of TaqMan gene expression assays used to examine levels of gene mRNA expression by qPCR.

Supplementary Table S4

RNA-seq statistics for the three cell lines treated with different siRNAs.

Supplementary Table S5

124 common DE genes between two *KMT2D* siRNAs vs. control comparisons.

Supplementary Table S6

31 common genes among all three *KMT2C* siRNA vs control comparisons.

Supplementary Table S7

Overrepresented biological pathways for genes with expression strongly correlated ($p < 0.001$) with that of *KMT2C/D* combined.

Supplementary Table S8

94 and 257 cell cycle genes significantly positively correlated with the combined *KMT2C/D* expression from the ICGC and TCGA datasets, respectively.

Supplementary Table S9

Three genes significantly positively correlated with the combined *KMT2C/D* expression, whose expression is also significantly downregulated in at least on *KMT2C/D* siRNA experiment.

Figure S1

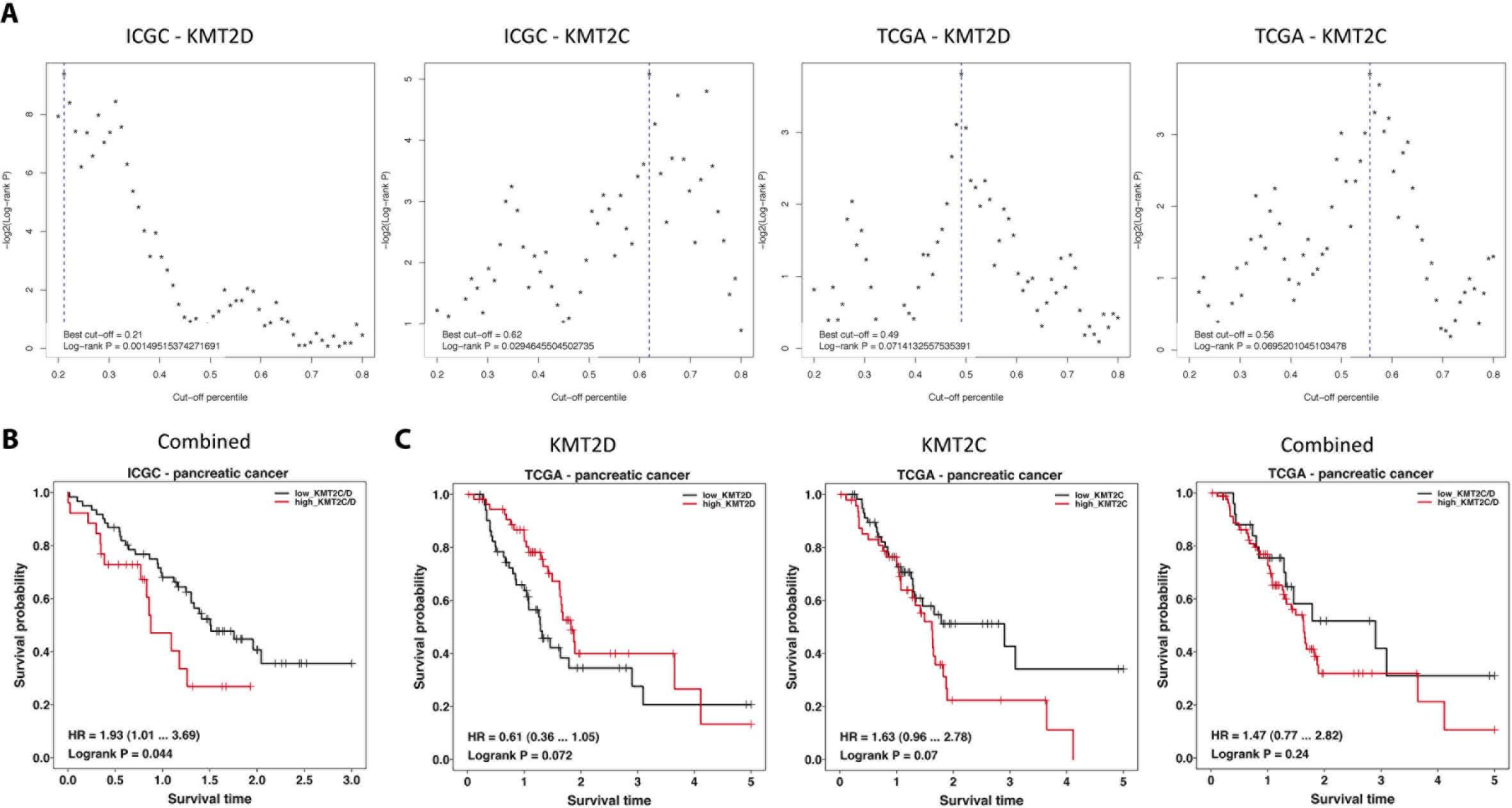

A
ICGC - KMT2D
Best cut−off = 0.21
Log−rank P = 0.00149515374271691
ICGC - KMT2C
Best cut−off = 0.62
Log−rank P = 0.0294645504502735
TCGA - KMT2D
Best cut−off = 0.49
Log−rank P = 0.0714132557535391
TCGA - KMT2C
Best cut−off = 0.56
Log−rank P = 0.0695201045103478
−log2(Log−rank P)
Cut−off percentile
B
Combined
ICGC - pancreatic cancer
low_KMT2C/D
high_KMT2C/D
HR = 1.93 (1.01 ... 3.69)
Logrank P = 0.044
C
KMT2D
TCGA - pancreatic cancer
low_KMT2D
high_KMT2D
HR = 0.61 (0.36 ... 1.05)
Logrank P = 0.072
KMT2C
TCGA - pancreatic cancer
low_KMT2C
high_KMT2C
HR = 1.63 (0.96 ... 2.78)
Logrank P = 0.07
Combined
TCGA - pancreatic cancer
low_KMT2C/D
high_KMT2C/D
HR = 1.47 (0.77 ... 2.82)
Logrank P = 0.24
Survival probability
Survival time

Figure S2

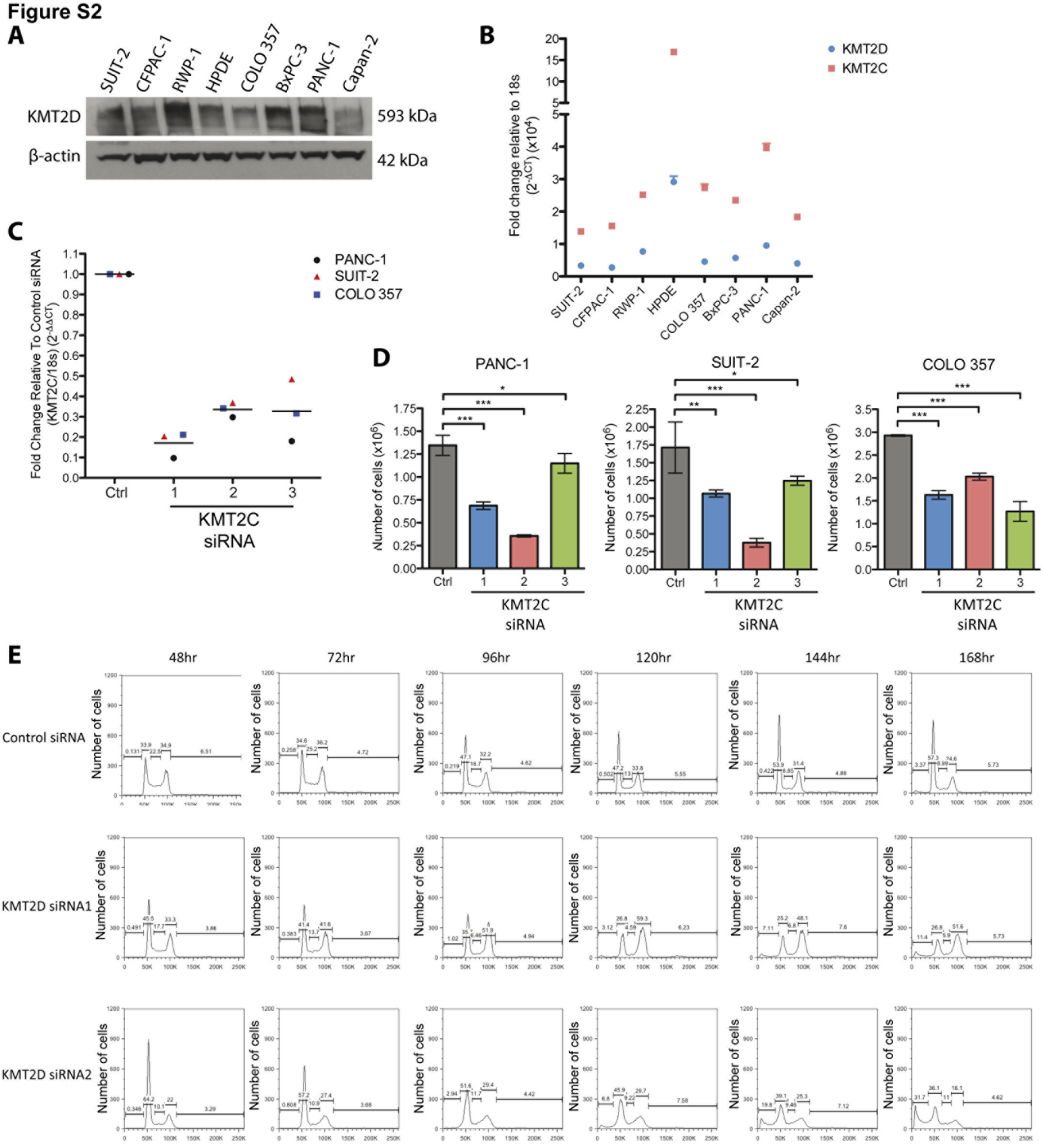

A
SUIT-2
CFPAC-1
RWP-1
HPDE
COLO 357
BxPC-3
PANC-1
Capan-2
KMT2D
593 kDa
β-actin
42 kDa
B
KMT2D
KMT2C
Fold change relative to 18s
C
PANC-1
SUIT-2
COLO 357
Fold Change Relative To Control siRNA
Ctrl
KMT2C siRNA
D
PANC-1
SUIT-2
COLO 357
Number of cells (x10^6)
E
48hr
72hr
96hr
120hr
144hr
168hr
Control siRNA
KMT2D siRNA1
KMT2D siRNA2
Number of cells

Figure S3

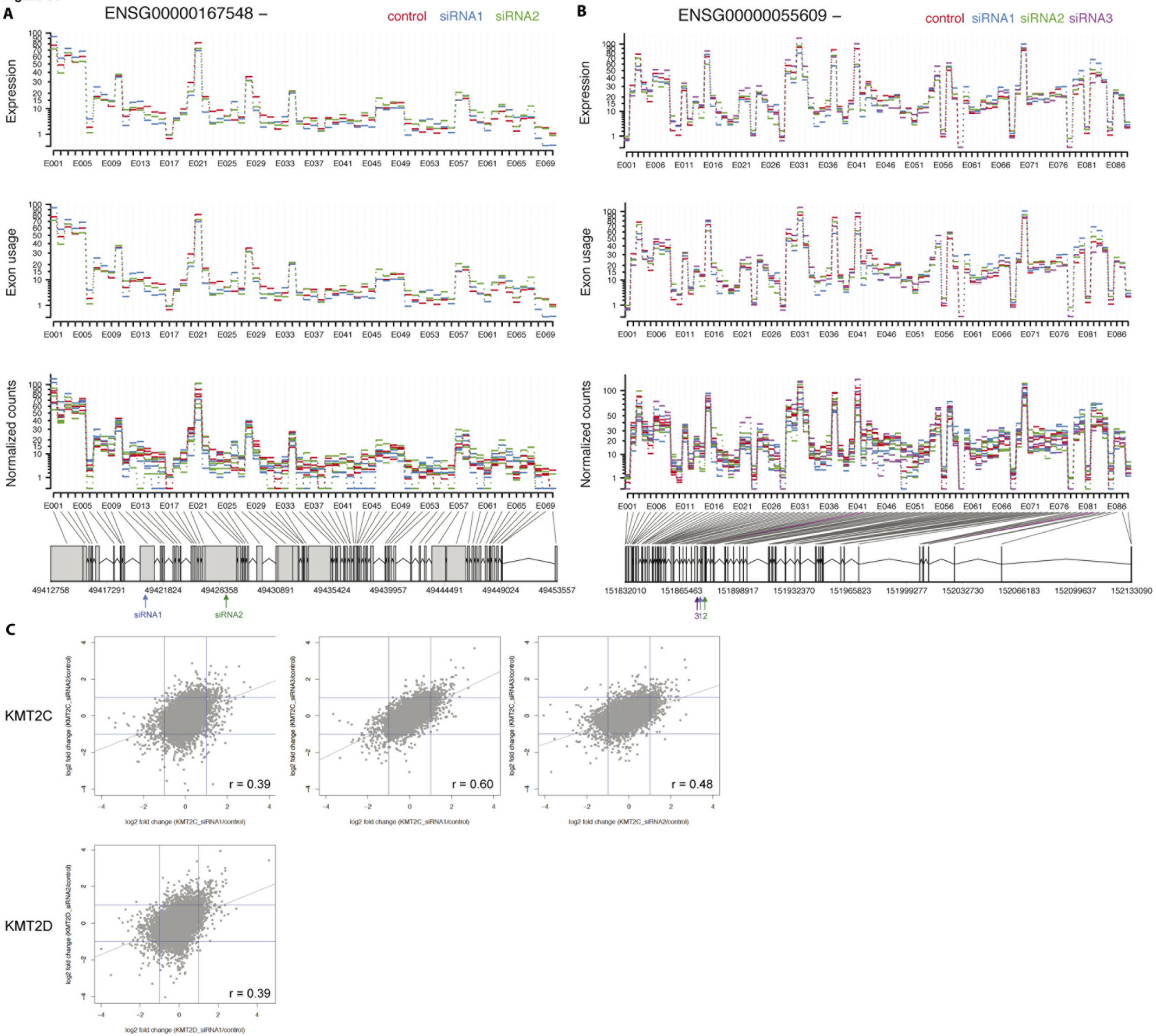

A
ENSG00000167548 –
control siRNA1 siRNA2
Expression
Exon usage
Normalized counts
E001 E005 E009 E013 E017 E021 E025 E029 E033 E037 E041 E045 E049 E053 E057 E061 E065 E069
49412758 49417291 49421824 49426358 49430891 49435424 49439957 49444491 49449024 49453557
siRNA1
siRNA2
B
ENSG00000055609 –
control siRNA1 siRNA2 siRNA3
Expression
Exon usage
Normalized counts
E001 E006 E011 E016 E021 E026 E031 E036 E041 E046 E051 E056 E061 E066 E071 E076 E081 E086
151832010 151865463 151898917 151932370 151965823 151999277 152032730 152066183 152099637 152133090
312
C
KMT2C
log2 fold change (KMT2C_siRNA2/control)
log2 fold change (KMT2C_siRNA1/control)
r = 0.39
log2 fold change (KMT2C_siRNA3/control)
log2 fold change (KMT2C_siRNA1/control)
r = 0.60
log2 fold change (KMT2C_siRNA3/control)
log2 fold change (KMT2C_siRNA2/control)
r = 0.48
KMT2D
log2 fold change (KMT2D_siRNA2/control)
log2 fold change (KMT2D_siRNA1/control)
r = 0.39

Figure S4

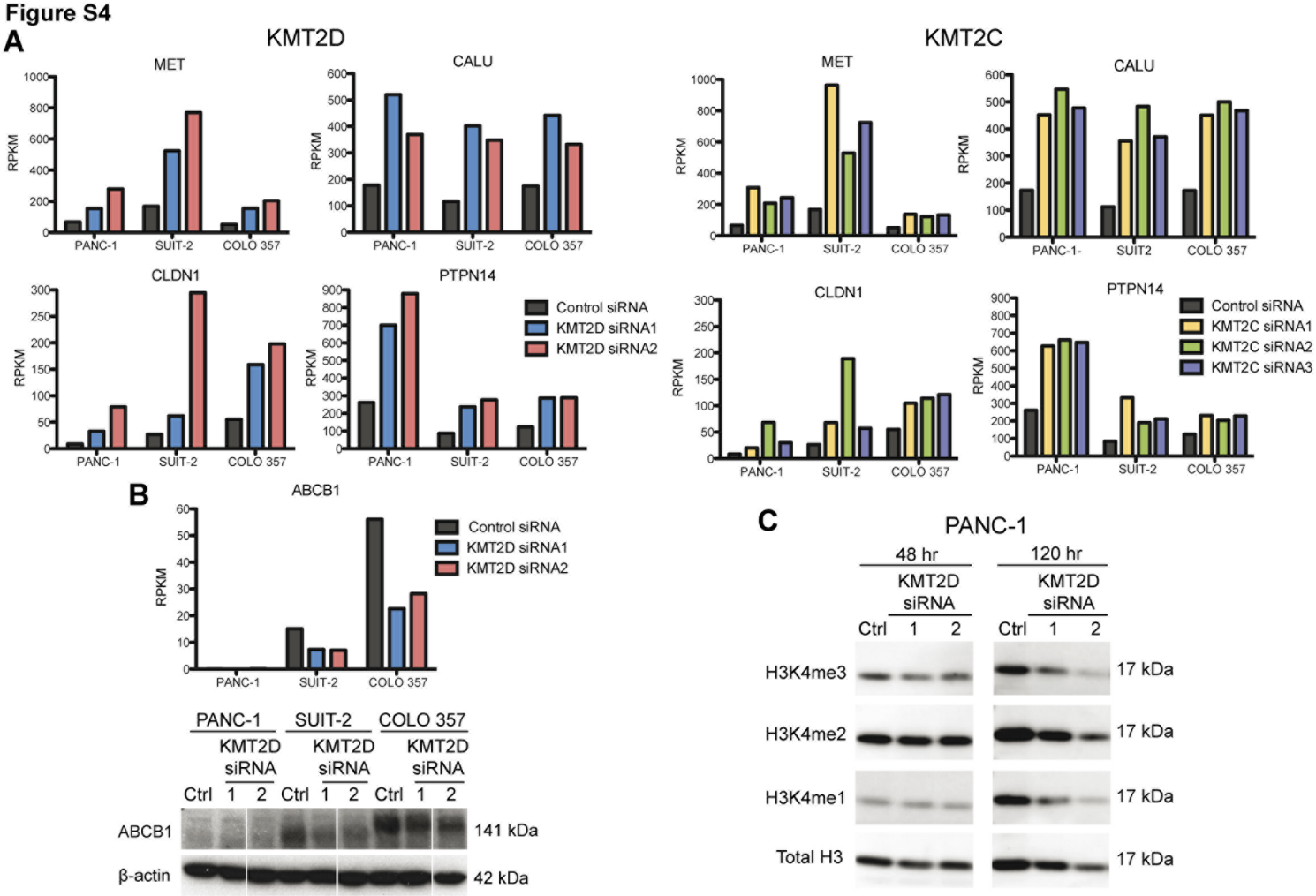

A
KMT2D
MET
CALU
CLDN1
PTPN14
RPKM
PANC-1
SUIT-2
COLO 357
Control siRNA
KMT2D siRNA1
KMT2D siRNA2
KMT2C
KMT2C siRNA1
KMT2C siRNA2
KMT2C siRNA3
B
ABCB1
PANC-1
SUIT-2
COLO 357
KMT2D siRNA
Ctrl
1
2
ABCB1
β-actin
141 kDa
42 kDa
C
PANC-1
48 hr
120 hr
H3K4me3
H3K4me2
H3K4me1
Total H3
17 kDa

Figure S5

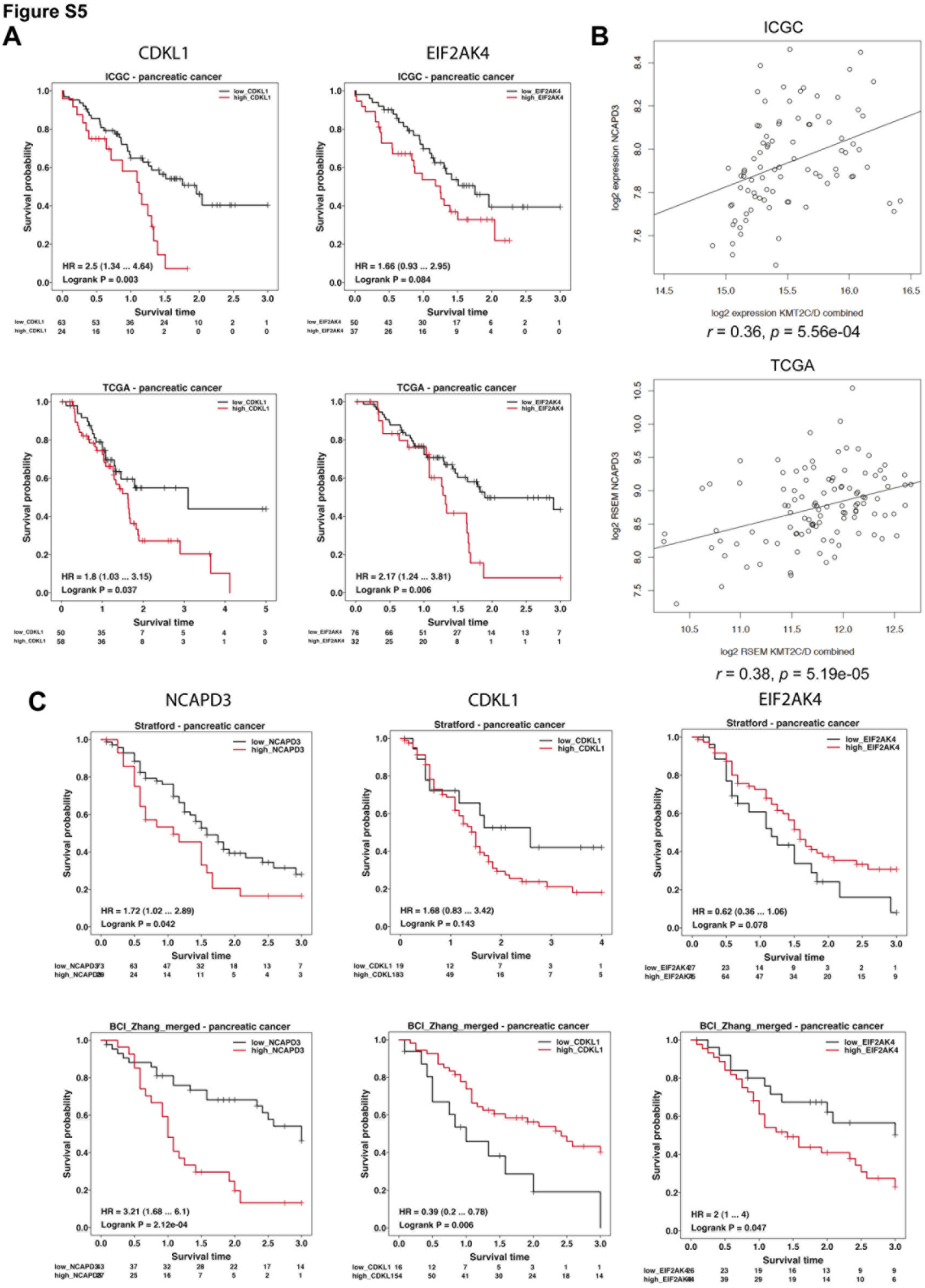
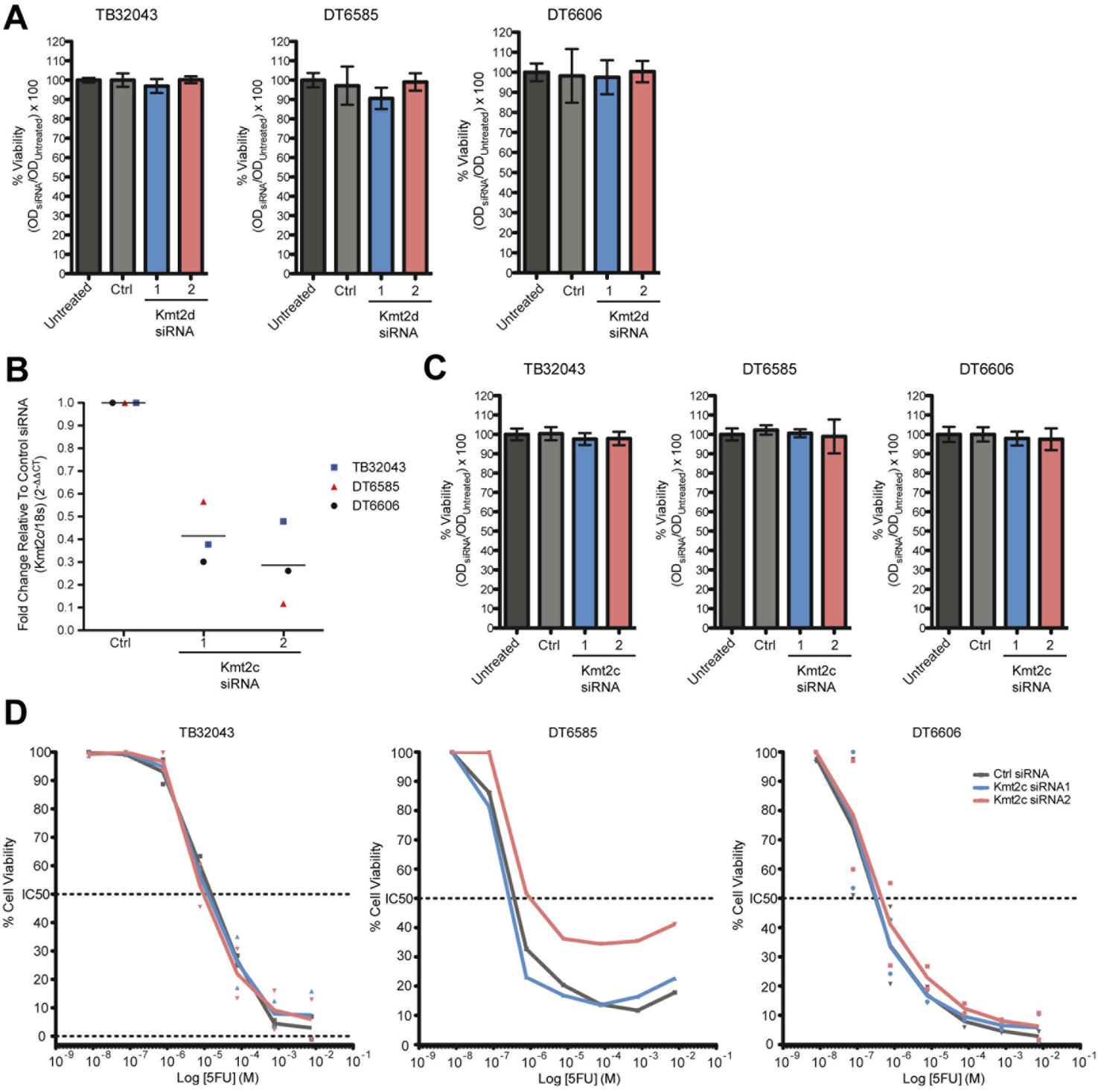

A
CDKL1
EIF2AK4
ICGC - pancreatic cancer
low_CDKL1
high_CDKL1
Survival probability
HR = 2.5 (1.34 ... 4.64)
Logrank P = 0.003
Survival time
low_CDKL1 63 53 36 24 10 2 1
high_CDKL1 24 16 10 2 0 0 0
low_EIF2AK4
high_EIF2AK4
HR = 1.66 (0.93 ... 2.95)
Logrank P = 0.084
low_EIF2AK4 50 43 30 17 6 2 1
high_EIF2AK4 37 26 16 9 4 0 0
TCGA - pancreatic cancer
HR = 1.8 (1.03 ... 3.15)
Logrank P = 0.037
low_CDKL1 50 35 7 5 4 3
high_CDKL1 58 36 8 3 1 0
HR = 2.17 (1.24 ... 3.81)
Logrank P = 0.006
low_EIF2AK4 76 66 51 27 14 13 7
high_EIF2AK4 32 25 20 8 1 1 1
B
ICGC
log2 expression NCAPD3
log2 expression KMT2C/D combined
r = 0.36, p = 5.56e-04
TCGA
log2 RSEM NCAPD3
log2 RSEM KMT2C/D combined
r = 0.38, p = 5.19e-05
C
NCAPD3
CDKL1
EIF2AK4
Stratford - pancreatic cancer
low_NCAPD3
high_NCAPD3
HR = 1.72 (1.02 ... 2.89)
Logrank P = 0.042
HR = 1.68 (0.83 ... 3.42)
Logrank P = 0.143
HR = 0.62 (0.36 ... 1.06)
Logrank P = 0.078
BCI_Zhang_merged - pancreatic cancer
HR = 3.21 (1.68 ... 6.1)
Logrank P = 2.12e-04
HR = 0.39 (0.2 ... 0.78)
Logrank P = 0.006
HR = 2 (1 ... 4)
Logrank P = 0.047

Figure S6


A
TB32043
DT6585
DT6606
% Viability
(OD siRNA/OD Untreated) x 100
Untreated
Ctrl
1
2
Kmt2d siRNA
B
Fold Change Relative To Control siRNA
(Kmt2c/18s) (2-ΔΔCT)
TB32043
DT6585
DT6606
Kmt2c siRNA
C
TB32043
DT6585
DT6606
Kmt2c siRNA
D
TB32043
DT6585
DT6606
% Cell Viability
IC50
Log [5FU] (M)
Ctrl siRNA
Kmt2c siRNA1
Kmt2c siRNA2