KMT2D Links TGF-β Signalling to Non-Canonical Activin Pathway and Regulates Pancreatic Cancer Cell Plasticity

Shuang Lu[1,2#], Hong Sun Kim[1#], Yubo Cao[1#], Karan Bedi[3], Lili Zhao[4], Ishwarya Venkata Narayanan[3], Brian Magnuson[4], Yumei Gu[1], Jing Yang[1], Zhujun Yi[1], Sepideh Babaniamansour[1], Sargis Shameon[1], Chang Xu[1], Michelle T. Paulsen[3], Ping Qiu[5], Sivakumar Jeyarajan[1], Mats Ljungman[3], Dafydd Thomas[1], Yali Dou[6], Howard Crawford[7], Marina Pasca di Magliano[8], Kai Ge[9], Bo Yang[5], Jiaqi Shi[1*]

Supplementary Information

Supplementary Materials and Methods

Bru-seq and BruChase-seq

Bru-seq and BruChase-seq were performed as previously described[1]. Briefly, cells were incubated in media containing bromouridine (Bru) (Sigma-Aldrich) at a final concentration of 2 mM for 30 minutes at 37°C to label nascent RNA. Cells were then lysed directly in Trizol, and total RNA was isolated. Bru-labeled RNA was immunocaptured using anti-BrdU antibodies, followed by the preparation of strand-specific cDNA libraries with the Illumina TruSeq kit (Illumina). The quality and concentration of the libraries were determined using an Agilent Bioanalyzer. Deep sequencing was performed at the University of Michigan Advanced Genomics Core using the Illumina HiSeq 2000 sequencer (51 bp read length, single-end sequencing) as previously described[1-3]. For BruChase-seq, cells were first labeled with 2 mM Bru for 30 minutes, washed in PBS, and then incubated in conditioned media containing 20 mM uridine for 6 hours. The cells were then lysed in Trizol, and Bru-labeled RNA was captured and processed as described above. Reads were pre-mapped (hence called premap fraction) to the ribosomal RNA (rRNA) repeating unit (GenBank U13369.1) and the mitochondrial and EBV genomes (from the hg38 analysis set) using Bowtie2 (2.3.3)[4]. Unaligned reads were subsequently mapped to human genome (hence called genome fraction) build hg38/GRCh38 using STAR (v 2.5.3a)[5] and a STAR index created from GENCODE annotation version 27 (ftp://ftp.ebi.ac.uk/pub/databases/gencode/Gencode human/release27/gencode.v27.basic.annotation.gtf.gz)[6]. The relative level of nascent RNA labeling was presented as RPKM values from Bru-seq data. Relative RNA stability is expressed as the ratio of BruChase-seq at 6-hour/ Bru-seq at 0-hour RPKM values. Synthesis was filtered for genes >300 bp and stability was filtered for 0-hour RPKM value>0.5 to ensure reliable numerical evaluations. The sequencing coverage and quality statistics for each sample are summarized in Supplementary Table 3.

Supplementary Table 1. miRNA and siRNA information

Hsp-miR-147b	GUGUGCGGAAAUGCUUCUGCU
KMT2D scramble siRNA	UGGUUUACAUGUCGACUAA
KMT2D siRNA1	GAAAGAAGCUGCGGAAGGA
KMT2D siRNA2	GCAGAUACCUUGUUUAGCA
KMT2D crRNA1	AGTGTGCCAAGCCTGCAGGT
KMT2D crRNA2	AGTGGCGGCACTGGATTAGT
ctrl siRNA	UGGUUUACAUGUCGACUAA
INHBA siRNA	GAACGGGUAUGUGGAGAUA

Supplementary Table 2. Information on antibodies and reagents

TGF-β1	Cell Signaling (#8915LC, 20ng/ml)
activin A	PEPROTECH (#120-14P, 10ng/ml)
follistatin	PEPROTECH (#120-13, 100ng/ml with activin A, 40ng/ml when using alone)
SB505124	Sigma-Aldrich (#S4696, 2.5uM)
SB202190	EMD Millipore (#559388, 5 or 10uM)
E-cadherin	Cell Signaling (#3195, 1/1000)
MMP2	Cell Signaling (#40994, 1/1000)
SNAIL	Cell Signaling (#3879, 1/1000)
GAPDH	Cell Signaling (#5174, 1/1000)
β-actin	Sigma-Aldrich (#A5441,1/2500)
Vimentin	Santa Cruz (#SC66002, 1/500)
H3K4me1	Cell Signaling (#5326s, 1/1000)
H3K4me2	EMD Millipore (#17677, 1/1000)
H3K4me3	Cell Signaling (#9751s, 1/1000)
H3	Cell Signaling (#4499, 1/1000)
KMT2D	EMD Millipore (#ABE1867, 1/1000)
p38	Cell Signaling (#8690, 1/1000)
p-p38	Cell Signaling (#4511, 1/1000)
ERK	Cell Signaling (#4691, 1/1000)
p-ERK	Cell Signaling (#2965, 1/1000)
Akt	Cell Signaling (#4691, 1/1000)
p-Akt	Cell Signaling (#2965, 1/1000)
JNK1	Cell Signaling (#3708, 1/1000)
p-SAPK/JNK	Cell Signaling (#4668, 1/1000)
SMAD2	Cell Signaling (#3103, 1/1000)
p-SMAD2	Cell Signaling (#3108s, 1/500)
SMAD3	Cell Signaling (#9523s, 1/1000)
p-SMAD3	Cell Signaling (#9520s, 1/500)

Supplementary Table 3. Summary of sequencing coverage and quality statistics

Sample ID	Total number of sequenced reads	Total number of uniquely mapped reads[a]	RNA integrity number (RIN)*	Ratio of all reads aligned to rRNA regions to total uniquely mapped reads (rRNA rate)**	Ratio of exon-mapped reads to total uniquely mapped reads (Expression Profile Efficiency)	Total number of detected transcripts with reads ⩾1[b]
KO1	71,191,915	57,483,106	N/A	0.079679704	0.135513712	6609
KO2	62,404,368	47,555,769	N/A	0.128660499	0.128217138	7547
WT	47,160,143	33,572,271	N/A	0.197165287	0.141446954	7697
Control/TGFb 0h	77,819,433	60,543,908	N/A	0.027509741	0.1104291	6228
Control (BruChase-seq)	92,669,081	44,061,771	N/A	0.063216787	0.520406336	6120
TGFb 1h	62,761,061	49,038,731	N/A	0.043510299	0.109346756	5806
TGFb 24h	74,032,398	58,047,522	N/A	0.023900055	0.110775045	5651
TGFb 72h	59,252,840	44,777,907	N/A	0.092399716	0.114080073	5735
72h TGFb (BruChase-seq)	41,714,886	25,847,419	N/A	0.042244402	0.496663578	5377

[a] GRCh38 reference genome was used.

[b] A higher minimum coverage threshold is permitted. We calculate RPKM values, instead of TPM values. In our case, these numbers represent the number of genes that meet a threshold of >= 1 RPKM and gene length >= 300 bp.

*We don’t analyze our total RNA (and therefore get RIN values) since we pull out the Bru-RNA from the total RNA and use that for library prep

** Total uniquely mapped reads = Uniquely mapped reads to premap fraction + uniquely mapped reads to genome fraction

Supplementary Table 4. Primer sequences

CDH1	forward 5'-GGCCTGAAGTGACTCGTAACG-3' reverse 5'-CAGTATCAGCCGCTTTCAGATTT-3'
MMP2	forward 5'-TTGATGGCATCGCTCAGATC-3' reverse 5'-TTGTCACGTGGCGTCACAGT-3'
SNAIL	forward 5'-CAGACCCACTCAGATGTCAA reverse 5'-CATAGTTAGTCACACCTCGT-3'
Vimentin	forward 5'-CCTTGAACGCAAAGTGGAATC-3' reverse 5'-GACATGCTGTTCCTGAATCTGAG-3'
KMT2D	forward 5'-GTCTTCACTGACGCCTCTCC-3' reverse 5'-GCTCCTCGCCCTTCAGATAC-3'
GAPDH	forward 5'-ACCCAGAAGACTGTGGAT-3' reverse 5'-GAGGCAGGGATGATGTTC-3'
INHBA	forward 5'-GTGAGTGCCCGAGCCATATAG-3' reverse 5'-CATGCGGTAGTGGTTGATGACT-3'

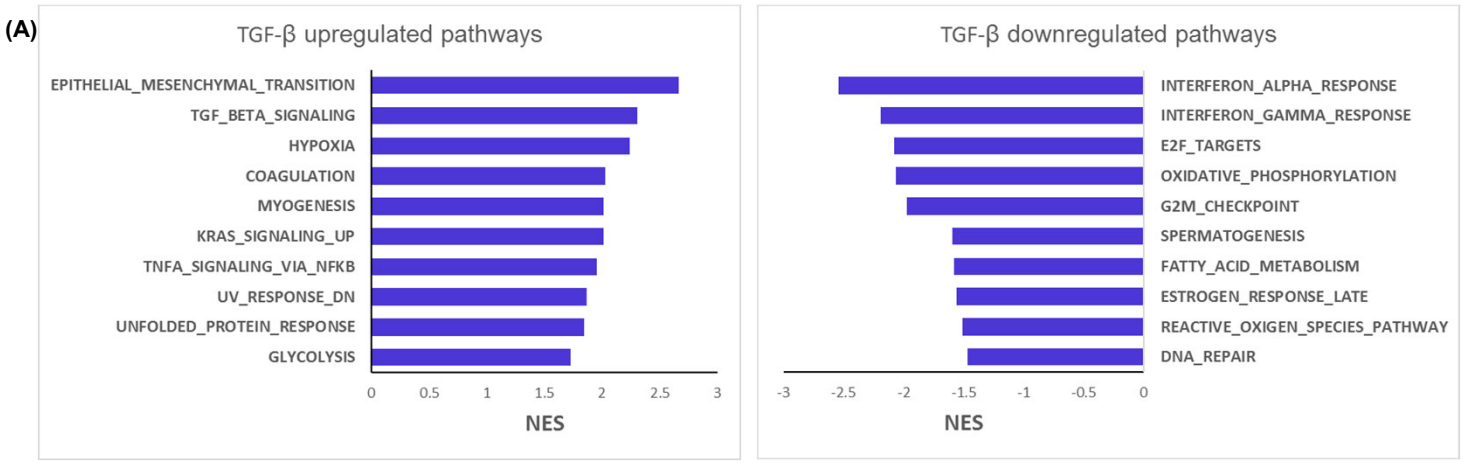
Supplementary Table 5. Protein expression levels of KMT2D in PDAC and its precursors

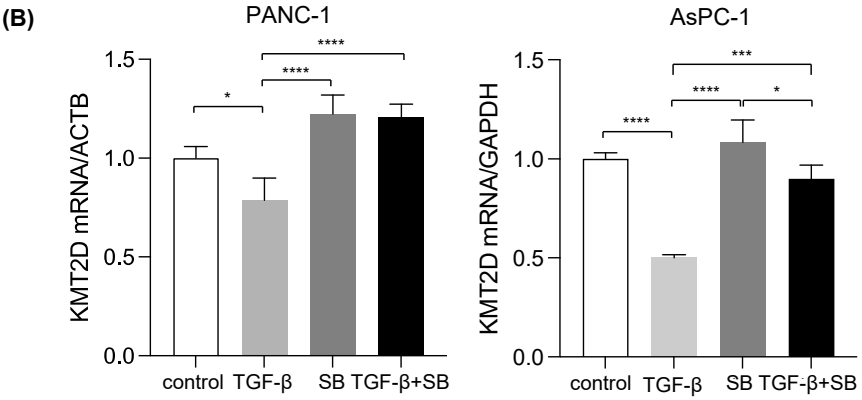
Category	N=	Mean ± SD	p-value	
Normal/Pancreatitis	35	192±46		
PanIN	30	194±45		
MCN	13	189±44		
IPMN	38	214±53		
PDAC	78	164±42	0.0000002	(vs. benign)
MET PDAC	19	147±37	0.0004	(vs. others)
Total	213			

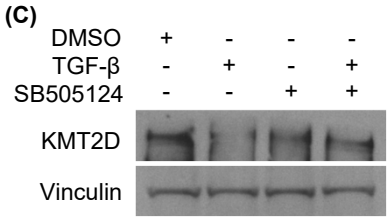
PanIN: Pancreatic intraepithelial neoplasia; IPMN: Intraductal papillary mucinous neoplasm; MCN: Mucinous cystic neoplasm; PDAC: Pancreatic ductal adenocarcinoma; MET PDAC: metastatic PDAC; SD: Standard deviation; vs.: versus.

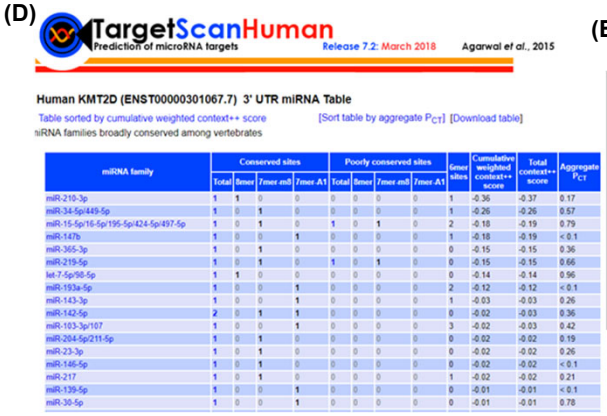
Supplementary Table 6. Genes upregulated >2 log2FC in KMT2D knockout BxPC3 cells (two clones, KO1 and KO2) compared to control

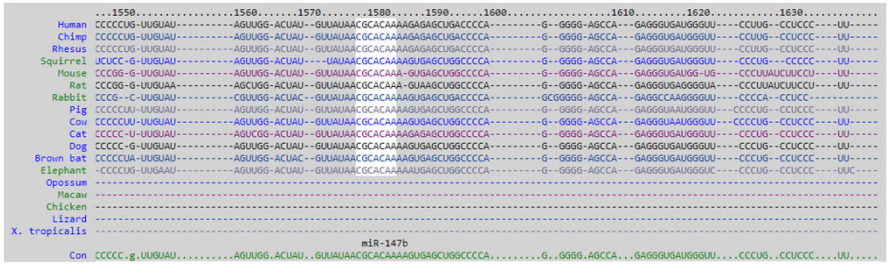
See attached excel file.

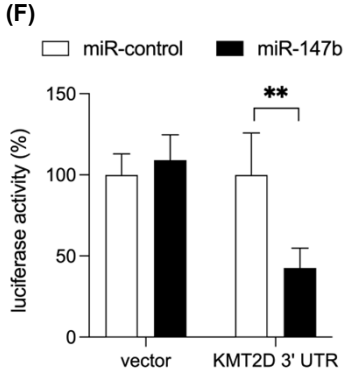


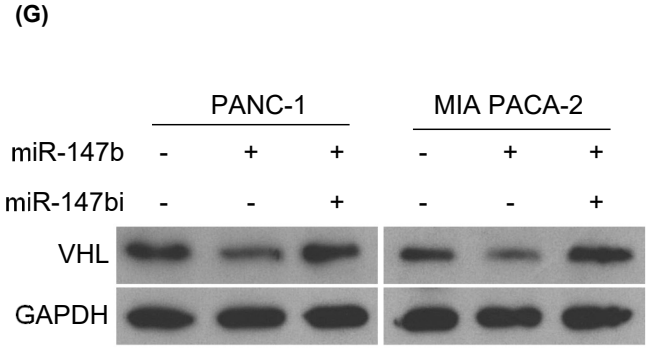


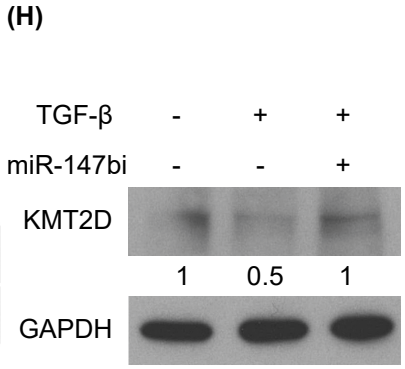






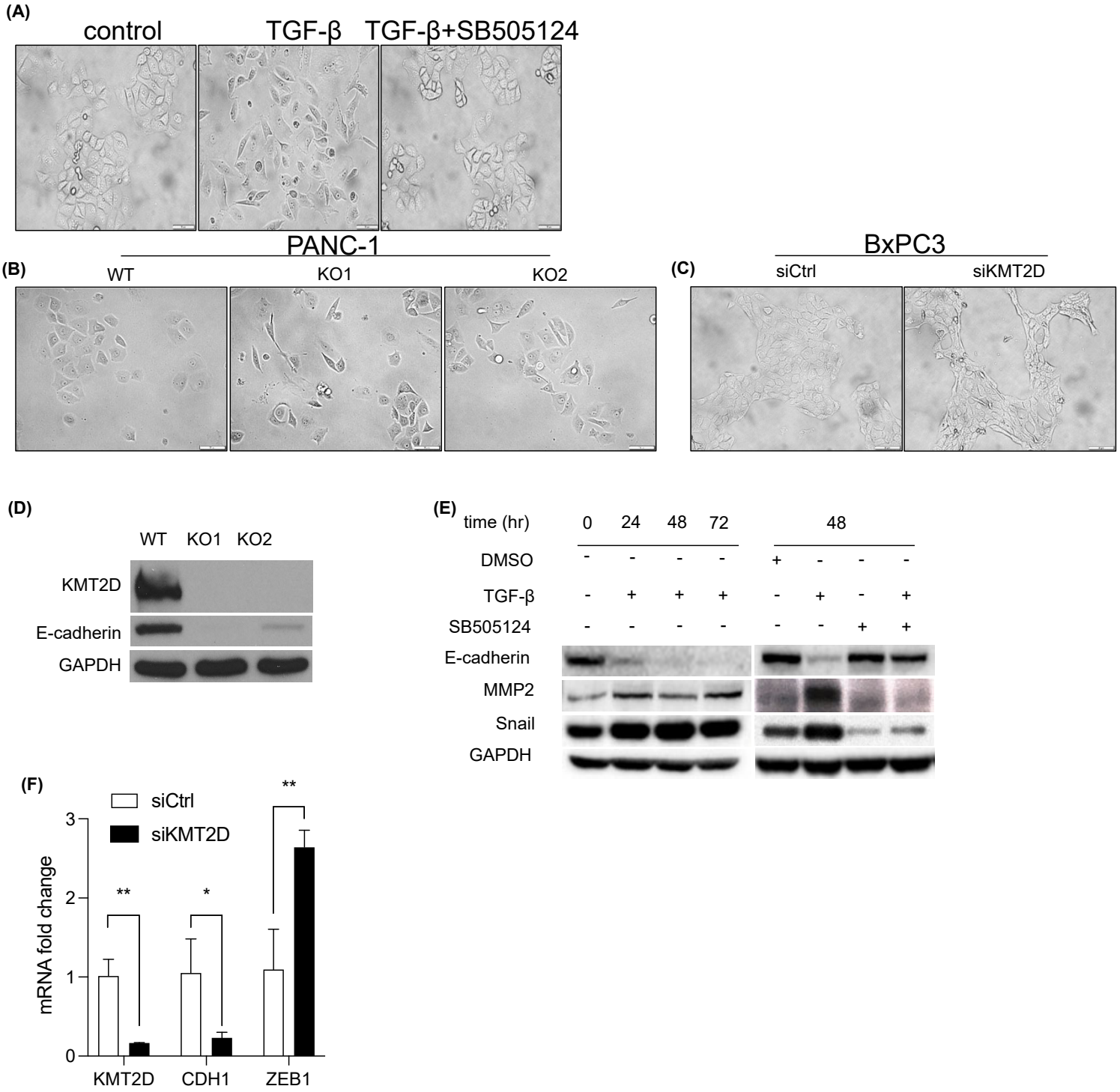






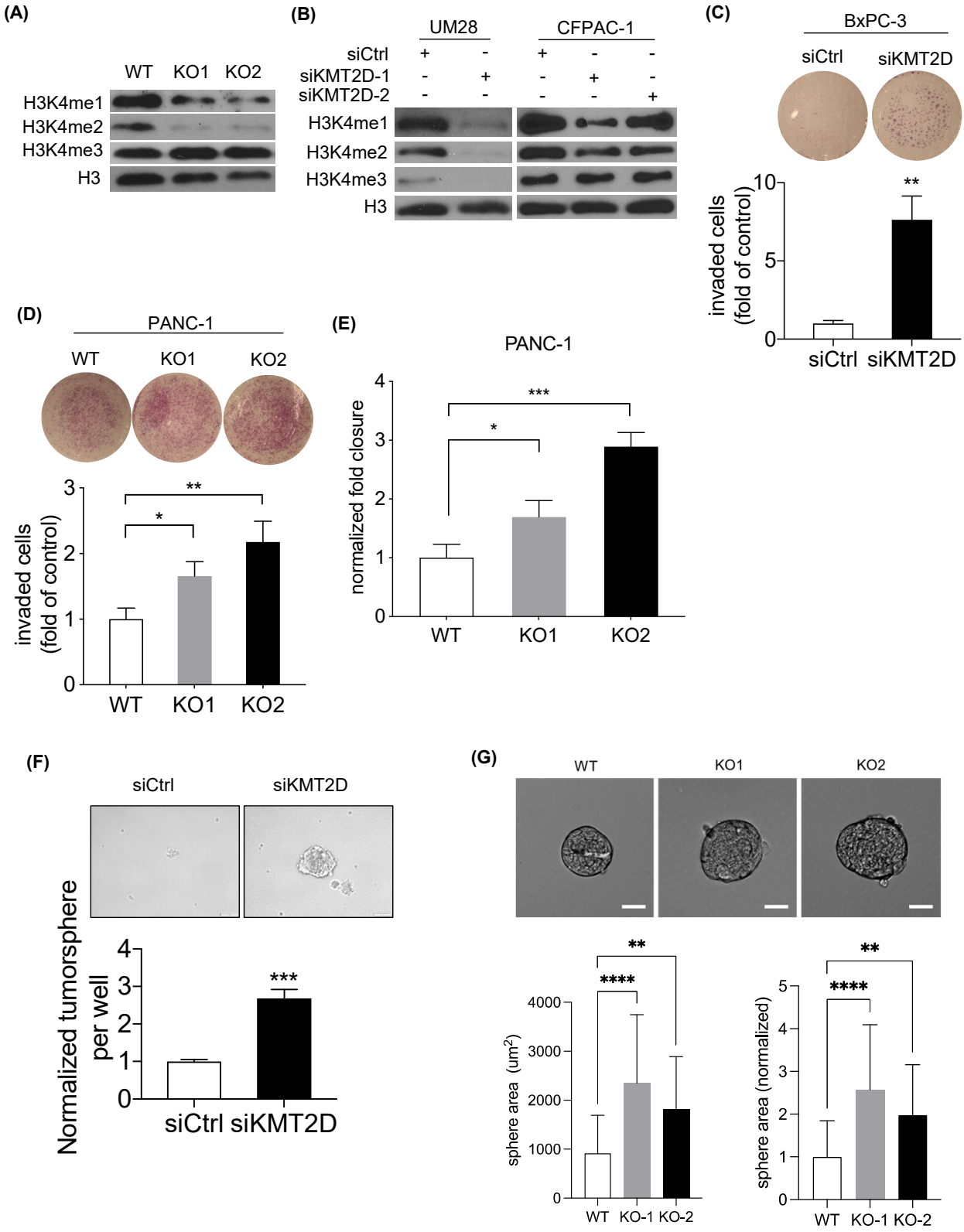
Supplementary Fig. 1 TGF-β downregulates KMT2D via miR-147b in PDAC cells. **(A)** Top hallmark pathways that are up- or down-regulated in TGF-β treated PANC-1 cells. NES: normalized enrichment score. **(B)** mRNA expression of KMT2D in PANC-1 and AsPC-1 cells treated with TGF-β with or without its inhibitor SB505124. **(C)** Western blot of MIA PaCa-2 cells treated with 20 ng/ml TGF-β with or without 2.5 nM SB505124 or DMSO for 48h. β-actin was used as loading control. **(D)** TargetScan prediction of potential miRNAs targeting 3' UTR of KMT2D. **(E)** Conserved site in miR-147b among mammals queried from TargetScanHuman. **(F)** Dual-luciferase reporter assay with negative control (miR-control) or miR-147b mimic in the presence or absence of predicted targeted KMT2D 3' UTR sequence. **(G)** Western blot of VHL after miR-147b mimic with or without miR-147bi transfection in two PDAC cell lines. **(H)** Western blot of KMT2D in the presence of TGF-β or TGF-β +miR-147bi. GAPDH was used as loading control. **p<0.01, ****p<0.001, One-way ANOVA test with Dunnett's multiple comparisons test (n=3).

Supplementary Figure 1



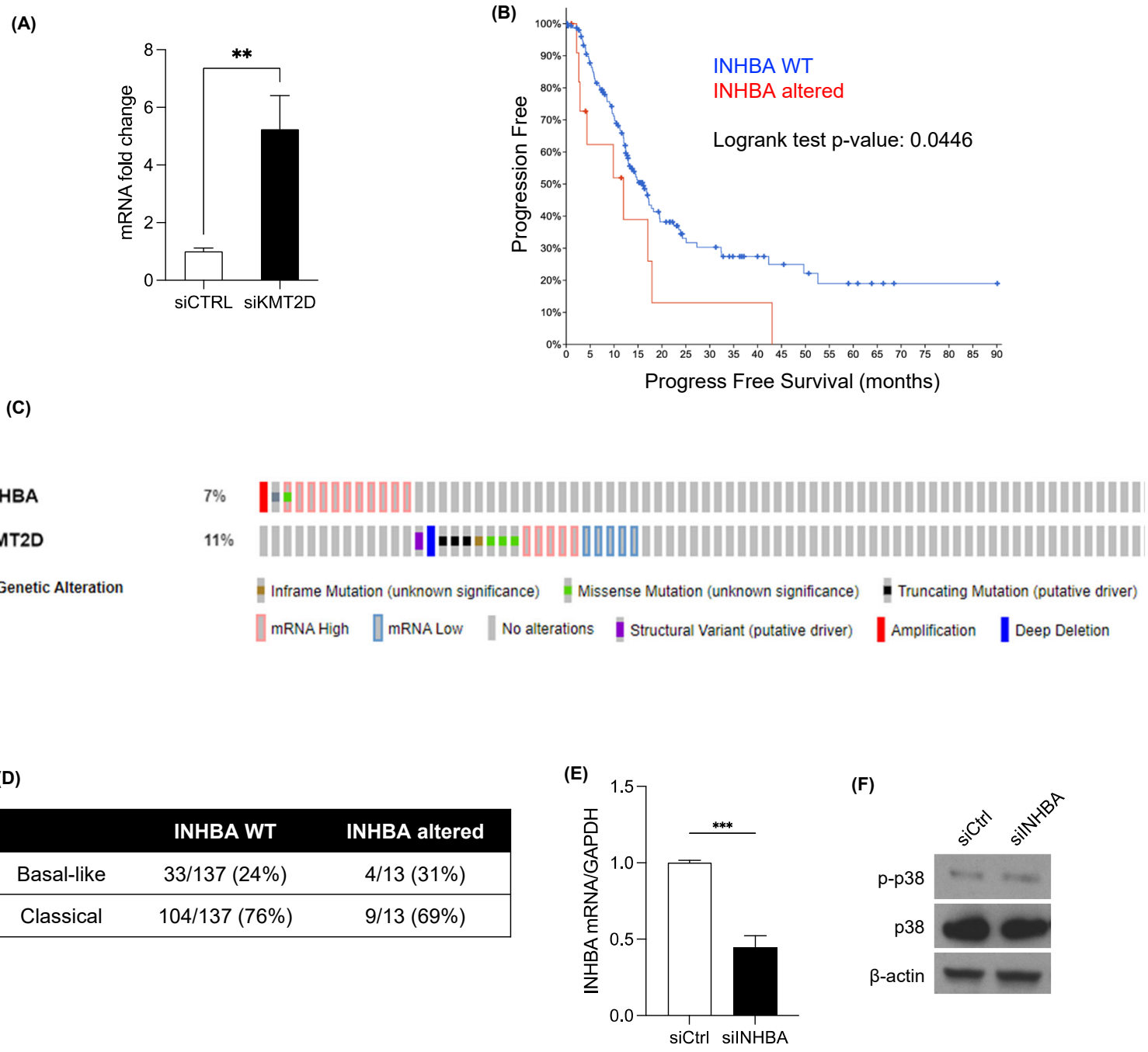
Supplementary Fig. 2 Loss of KMT2D induces EMT in PDAC cells. **(A)** Phase-contrast images of PANC-1 cells treated with DMSO, or TGF-β with or without SB505124. (scale bar = 100 μm). **(B)** Phase-contrast images of PANC-1 wild-type (WT) cells and KMT2D knockout (KO) cells. Scale bar: 100 μm. **(C)** Phase-contrast images of BxPC-3 cells transfected with 50 nM scrambled siRNA (siCtrl) or KMT2D siRNA (siKMT2D) for 5 days. **(D)** Western blot analysis of KMT2D and E-cadherin in WT and KMT2D-KO PANC-1 cells. GAPDH was used as loading control. **(E)** Western blot analysis of E-cadherin, MMP2, and Snail in PANC-1 cells treated with DMSO or 20 ng/ml TGF-β with or without SB505124 for 24, 48, and 72 hours. GAPDH was used as loading control. **(F)** Real-time RT-PCR of KMT2D, CDH1, and ZEB1 in control and KMT2D siRNA knockdown HPDE cells (*p<0.05, **p<0.01, multiple unpaired student t test, n=3)

Supplementary Figure 2



Supplementary Fig. 3 Reduced KMT2D expression decreases global H3K4 mono- and di-methylation and promotes cell migration, invasion, and tumorigenicity in PDAC cells. **(A)** Western blot of histone H3 lysine4 methylation levels in WT and KMT2D-KO BxPC-3 cells. H3 was used as loading control. **(B)** Western blot analysis of histone H3 lysine4 methylation levels in siCtrl or siKMT2D transfected UM28 or CFPAC-1 PDAC cell lines. H3 was used as loading control. **(C)** Invaded BxPC-3 cells transfected with either siCtrl or siKMT2D at 48h by transwell invasion assay (**p<0.01, unpaired student t test, n=3). **(D)** Invaded cells in 2 KMT2D-KO PANC-1 clones and WT PANC-1 cells at 48 h by transwell invasion assay. (*p<0.05, **p<0.01, one-way ANOVA with Dunnett's multiple comparisons test, n=3) **(E)** Wound closure in KMT2D-KO and WT PANC-1 cells at 48h post scratch by wound healing assay. **(F)** Tumor spheres formed by siCtrl or siKMT2D transfected UM28 cells by tumor sphere formation assay. **(G)** Tumor sphere size formed by control or KMT2D-knockout BxPC-3 cells by tumor sphere formation assay. Scale bar=25 μm. **p<0.01, ***p<0.005, ****p<0.001, unpaired student t test, n=3

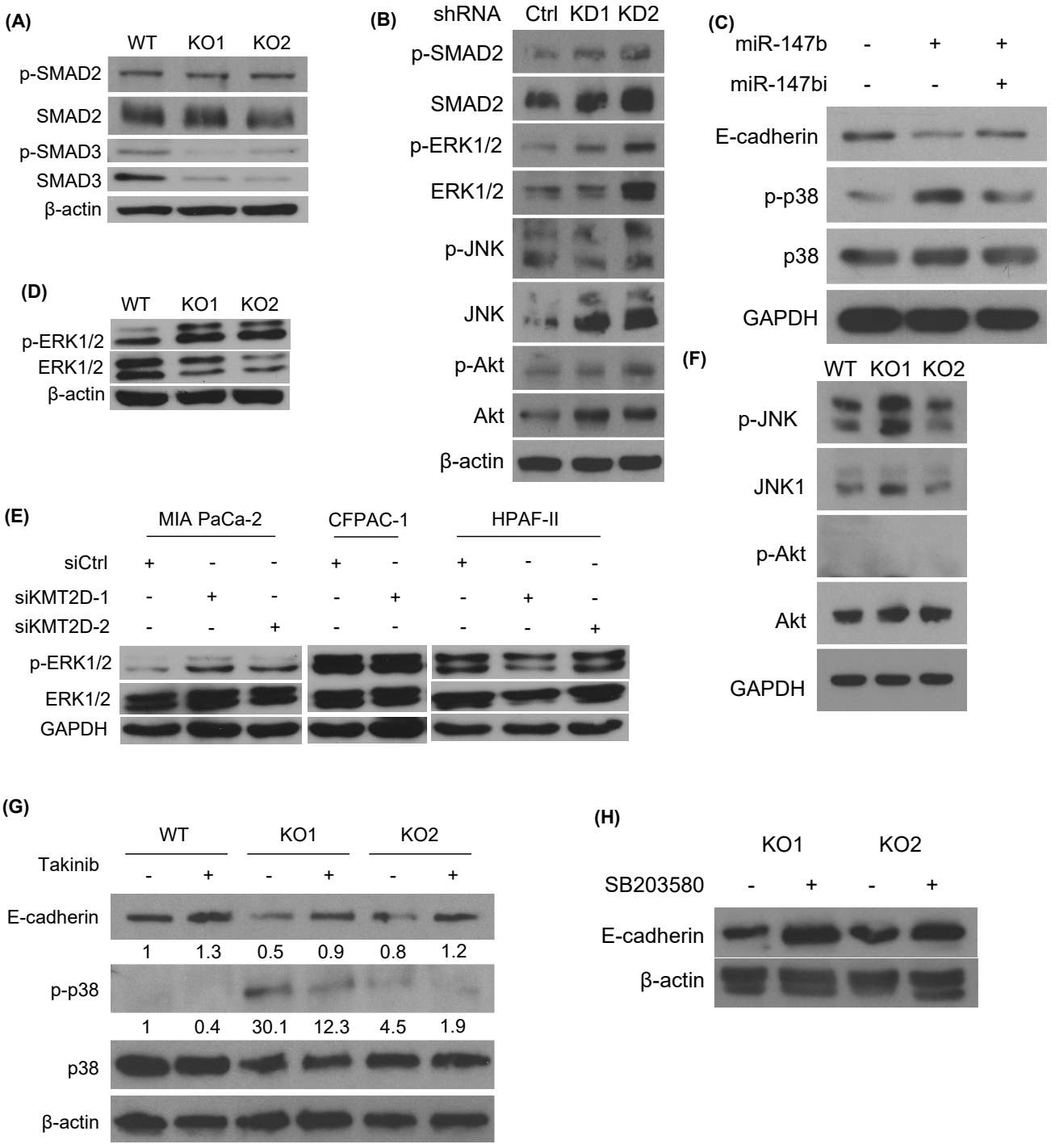
Supplementary Figure 3



	INHBA WT	INHBA altered
Basal-like	33/137 (24%)	4/13 (31%)
Classical	104/137 (76%)	9/13 (69%)

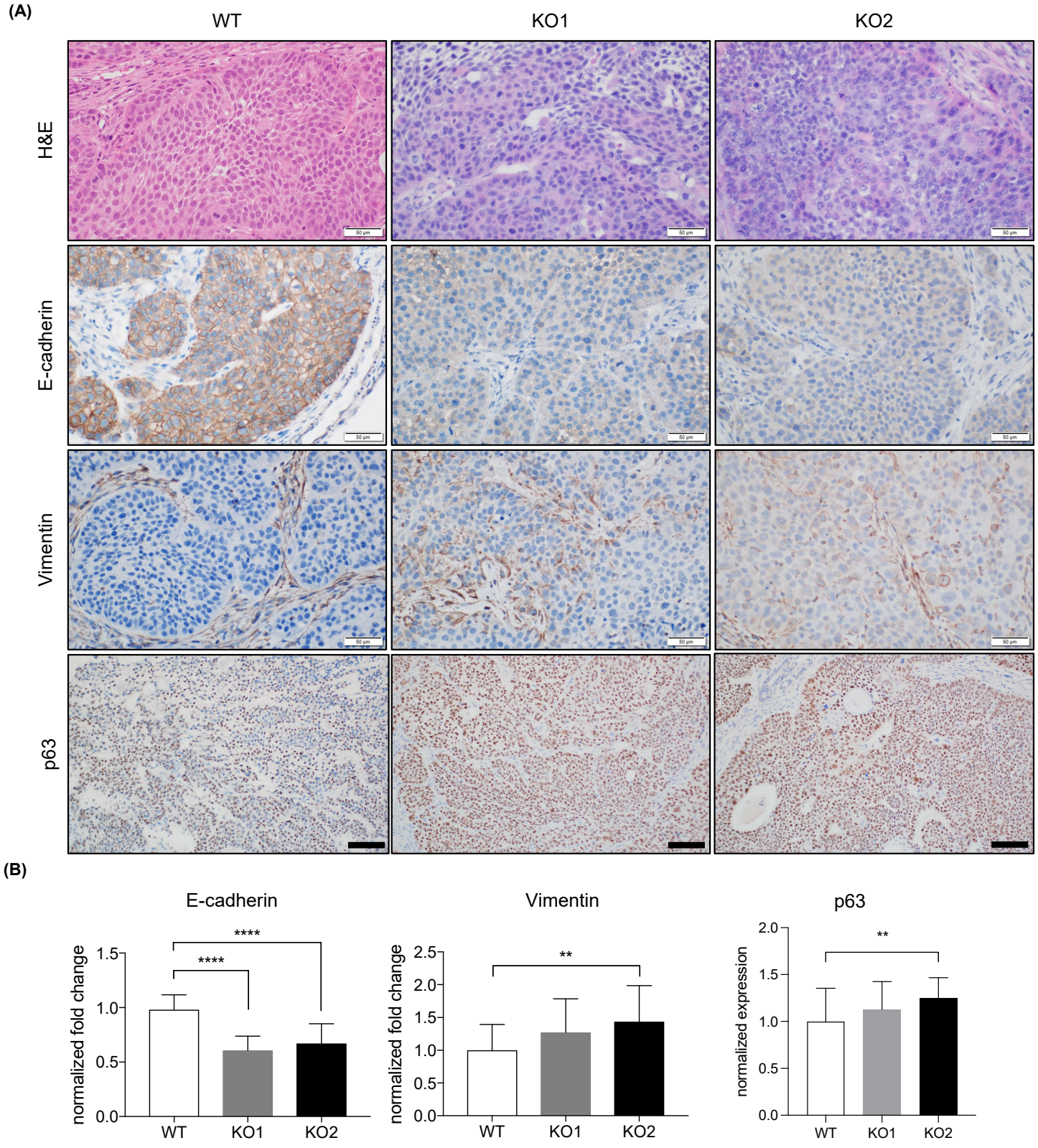
Supplementary Fig. 4 KMT2D depletion induces activin A expression. **(A)** *INHBA* mRNA expression quantified by RT-qPCR in HPDE cells transfected with siKMT2D or siCTRL. (**p<0.01, unpaired student t test, n=3) **(B)** Kaplan-Meir curve of progression-free survival in PDAC patients with or without INHBA alterations (n=168, TCGA database). **(C)** Genetic and transcriptional alterations of INHBA and KMT2D in human PDAC tissues (n=168, TCGA database). **(D)** Percentage of basal-like and classical PDAC subtypes in INHBA WT and altered tumors. **(E)** Real-time RT-PCR of *INHBA* in control (siCtrl) or *INHBA* knockdown (siINHBA) transfected BxPC-3 WT cells. GAPDH was used as the reference gene (***p<0.0005, unpaired student t-test, n = 3). **(F)** Western blot of p-p38 and p38 in WT BxPC-3 cells transfected with siCtrl or siINHBA for 24 hours. β-actin was used as loading control.

Supplementary Figure 4



Supplementary Fig. 5 miR-147b induced KMT2D loss activates a non-canonical activin A pathway. **(A)** Western blot analysis of p-SMAD2, SMAD2, p-SMAD3, and SMAD3 in KMT2D WT and KO BxPC-3 cells. Same β-actin loading control was shown as in Fig. 5f. **(B)** Western blot analysis of p-SMAD2, SAMD2, p-ERK1/2, ERK1/2, p-JNK, JNK, p-Akt, and Akt in WT and KMT2D-KD PANC-1 cells. β-actin was used as loading control. **(C)** Western blot of E-cadherin, p-p38, and p38 in PANC-1 cells transfected with miR-147b mimic with or without miR-147bi. GAPDH was used as loading control. **(D)** Western blot analysis of ERK1/2 and p-ERK1/2 in WT and KMT2D-KO BxPC-3 cells. β-actin was used as loading control. **(E)** Western blot analysis of ERK1/2 and p-ERK1/2 in PDAC cells transfected with siCtrl or siKMT2D. Same GAPDH loading control for MIA PaCa-2 and HPAF-II was shown as in Fig. 5g. **(F)** Western blot analysis of p-JNK, JNK, p-Akt, and Akt in WT and KMT2D-KO BxPC-3 cells. GAPDH was used as loading control. **(G)** Western blot of E-cadherin, p-p38, and p38 in KMT2D-KO BxPC-3 cells treated with either control or TAK1 inhibitor, Takinib, or DMSO for 72 hours. β-actin was used as loading control. Band quantification for E-cadherin/B-actin and p-p38/p38 are listed. Quantification was normalized to WT DMSO. **(H)** Western blot of E-cadherin in KMT2D-KO BxPC-3 cells after 96 hours treatment with SB203580. β-actin was used as loading control.

Supplementary Figure 5



Supplementary Fig. 6 KMT2D knockout PDAC cells form tumors with enhanced mesenchymal differentiation in vivo. **(A)** Representative H&E and immunohistochemical stains of E-cadherin, Vimentin, and p63 in BxPC-3 WT and KMT2D-KO orthotopic xenograft tumors. (scale bar = 50 μm, p63 scale bar = 100 μm) **(B)** Quantification of immunohistochemistry stains of E-cadherin, Vimentin, and p63 in WT and KMT2D KO xenograft tumors shown in **a** (**p<0.01, ****p<0.001, One-way ANOVA test with Tukey's multiple comparisons, n=10)

Supplementary Figure 6

Supplementary References:

1. Paulsen, M.T., Veloso, A., Prasad, J., Bedi, K., Ljungman, E.A., Magnuson, B., Wilson, T.E. & Ljungman, M. Use of Bru-Seq and BruChase-Seq for genome-wide assessment of the synthesis and stability of RNA. *Methods.* 2014; **67**(1): 45-54.
2. Bedi, K., Magnuson, B.R., Narayanan, I., Paulsen, M., Wilson, T.E. & Ljungman, M. Co-transcriptional splicing efficiencies differ within genes and between cell types. *RNA.* 2021.
3. Paulsen, M.T., Veloso, A., Prasad, J., Bedi, K., Ljungman, E.A., Tsan, Y.C., Chang, C.W., Tarrier, B., Washburn, J.G., Lyons, R., Robinson, D.R., Kumar-Sinha, C., Wilson, T.E. & Ljungman, M. Coordinated regulation of synthesis and stability of RNA during the acute TNF-induced proinflammatory response. *Proc Natl Acad Sci U S A.* 2013; **110**(6): 2240-2245.
4. Langmead, B. & Salzberg, S.L. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* 2012; **9**(4): 357-359.
5. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. & Gingeras, T.R. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics.* 2013; **29**(1): 15-21.
6. Frankish, A., Diekhans, M., Ferreira, A.M., Johnson, R., Jungreis, I., Loveland, J., Mudge, J.M., Sisu, C., Wright, J., Armstrong, J., Barnes, I., Berry, A., Bignell, A., Carbonell Sala, S., Chrast, J., Cunningham, F., Di Domenico, T., Donaldson, S., Fiddes, I.T., Garcia Giron, C., Gonzalez, J.M., Grego, T., Hardy, M., Hourlier, T., Hunt, T., Izuogu, O.G., Lagarde, J., Martin, F.J., Martinez, L., Mohanan, S., Muir, P., Navarro, F.C.P., Parker, A., Pei, B., Pozo, F., Ruffier, M., Schmitt, B.M., Stapleton, E., Suner, M.M., Sycheva, I., Uszczynska-Ratajczak, B., Xu, J., Yates, A., Zerbino, D., Zhang, Y., Aken, B., Choudhary, J.S., Gerstein, M., Guigo, R., Hubbard, T.J.P., Kellis, M., Paten, B., Reymond, A., Tress, M.L. & Flicek, P. GENCODE reference annotation for the human and mouse genomes. *Nucleic Acids Res.* 2019; **47**(D1): D766-D773.