Supplemental Methods

Genetic knockins, knockouts, knockdowns, and overexpression

CRISPR-mediated knockouts were generated by cloning sgRNAs into the lentiGuide-Puro (Addgene plasmid # 52963) or pSpCas9(BB)-2A-Puro (PX459) (Addgene plasmid # 48139) vectors as previously described [1].

For CRISPR-mediated knock-in of the GFP tag to the 3' end of the endogenous ID1 locus, *Id1*-cterm sgRNA was cloned into the pSpCas9(BB)-2A-Puro vector, and HA-L, GFP, HA-R were cloned into the pUC19 vector using the KpnI, BamHI, EcoRI, and XbaI restriction sites. The donor plasmid was cleaved at KpnI and XbaI, and the fragment containing the tag and flanking homology arms was isolated by gel purification.

shRNA-mediated knockdowns were performed with mir-E-based shRNAs cloned into the doxycycline-inducible lentiviral LT3GEP vector. Except for tumor and spheroid formation assays with ID protein depletion, all shRNAs were induced for 3 days in doxycycline before the beginning of the assay. To maintain cell viability, expression of shRNAs targeting *Id* mRNAs was induced at the start of the experiment.

The pLVX-Tight-Puro vector (Clontech) was used for doxycycline-inducible cDNA expression, and the pLVX-IRES-Hyg vector (Clontech) was used for constitutive expression. *Id1* was cloned into the pLVX-Tight-Puro site using the BamHI/EcoRI sites. Doxycycline (1 μg/mL unless otherwise noted, Fisher) was added 12 h prior to experiments to induce gene expression.

For the CRISPR-mediated knock-in of the GFP tag to the 3' end of the endogenous *Id1* locus, 1.25 μg of the Cas9 vector and 1.25 μg of the cleaved donor fragment were transiently tranfected using Lipofectamine 3000 in 6-well plates seeded overnight with 10^5 cells. Cells were selected in 10 μg/mL puromycin (Sigma) for 48 h, recovered for 48 h, and GFP-positive clones were identified by FACS.

Immunoblotting

After transfer, membranes were incubated overnight in designated antibody, washed 3x in PBST, and incubated 1 h in secondary antibody in PBST plus 5% milk. Detection using the 680 and 800 channels of the Odyssey CLx imager. Primary antibodies were used at 1:1000 dilution except for anti-Id1 antibodies, which were used at 1:2500.

Immunohistochemistry (IHC) and Immunofluorescence (IF)

Tissue blocks were prepared from tissues fixed in 4% paraformaldehyde overnight and dehydrated in 70% alcohol. Paraffin-embedded sections were rehydrated using Histo-Clear (National Diagnostics) followed by 100-70% ethanol. Endogenous peroxidase activity was quenched with H_2O_2, antigen retrieval was performed in a steamer for 30 min in citrate antigen retrieval solution. Cell cultures plated on glass coverslips were fixed for 10 min a room temperature in 3% paraformaldehyde. For IHC, avidin and biotin blocking solutions (Vector Labs) were followed by normal horse serum (Vector Labs) and overnight incubation in the designated primary antibody. ImmPRESS HRP Anti-Rabbit Ig and ImmPACT DAB Peroxidase (Vector Labs) were used for detection. Detection was followed by dehydration of tissue in 70-100% ethanol and HistoClear, followed by mounting with Vectashield mounting medium. For IF, permeabilization in PBS plus 0.1% Triton X (PBST) was performed, followed by a 1 h block in PBSTr supplemented with 1% normal goat serum and 1% bovine serum albumin. Primary antibody and secondary antibody incubation of 45 min is separated by 3 washes with PBST and finished in 3 washes of PBST, 5 min incubation in 1x Hoechst, and mounting with Invitrogen ProLong Gold. Primary antibodies were used at 1:100 except for anti-Id antibodies, which were used at 1:500.

Flow cytometry

Kras^{G12D};Cdkn2a$^{-/-}$;Smad4$^{-/-}$ orthotopic tumors were collected 6 weeks post-implantation. Tumors were cut into 1 mm cubes and incubated for 30 min at 37ºC in 2 mg/mL collagenase V (Sigma, C9263) supplemented with 50 μM $CaCl_2$ and 5 mM glucose. The solution was filtered through a 70 μm filter and the pieces remaining in the filter were digested in TrypLE (Gibco) for 10 min at 37ºC. The resulting single-cell suspension was filtered through 70 μm filters (Cat. 542070, Greiner Bio-One) and centrifuged for 5 min at 1500 rpm. Cell pellets were resuspended in FACS Buffer (PBS supplemented with 0.5% bovine serum albumin and 2mM EDTA). To stain for cell surface markers of pancreatic cancer stem cells, the following antibodies were used: anti-CXCR4-PE (1:100, Cat. 146505, BioLegend), anti-SSEA-4-FITC (1:100, Cat. 330409, BioLegend), anti-CD44-PE (1:100, Cat. 12-0441-83, eBioscience) and anti-CD24-FITC (1:100, Cat. 11-0241-81, eBioscience). Anti-CD45-BUV395 (1:200, Cat. 565967, BD Biosciences) was used to exclude immune cells and Zombie Aqua (1:200, Cat. 423101, BioLegend) was used for exclusion of dead cells. Cells were blocked with anti-mouse CD16/32 (Cat. 14-0161-82, eBioscience) for 15 min at 4ºC, incubated with primary antibodies and Zombie Aqua for 30 min

at 4 °C and washed with FACS buffer. To co-stain with ID1, rabbit anti-mouse ID1 (1:250, Cat. BCH-1/37-2-100, Biocheck) was used. Rat anti-mouse CK19 (1:100, DSHB TROMA-III) was applied to stain for cancer cells. After cell surface staining, cells were fixed and permeabilized with Foxp3/Transcription Factor Staining Buffer Set (Cat. 00-5523-00) for intracellular staining of ID1 and CK19, following instructions from the manufacturer. Donkey anti-rabbit-Alexa Fluor 647 and Donkey anti-rat-Alexa Fluor 594 secondary antibodies were used to analyze expression of ID1 and CK19. Data were acquired on BD Fortessa and analyzed with FlowJo.

Cell viability and apoptosis

CellTiter-Glo and Caspase-Glo 3/7 (Promega) was used according to manufacturer's protocol with incubation for 1 h. Caspase 3/7 signal was normalized to CellTiter-Glo signal to account for increased apoptosis in larger cell populations.

RNA-seq and qRT-PCR

Total RNA was purified using Qiagen RNeasy Mini Kit. For RNA-seq, RNA quality and quantity were checked by Agilent BioAnalyzer 2000. 500 ng RNA per sample with RIN > 9.5 was used for library construction with TruSeq RNA Sample Prep Kit v2 (Illumina) according to manufacturer's instructions. Libraries were multiplex sequenced on a Hiseq2500 platform, and more than 20 million raw paired-end reads were generated for each sample. cDNA was synthesized from 100-1000 ng RNA using random hexamer oligos and the Transcriptor First Strand cDNA Synthesis Kit (Roche). qPCR was performed on a ViiA 7 Real-time PCR System (Applied Biosystems). Values are mean±range, representative of at least 2 independent experiments.

Genome sequencing

All organoid lines were sequenced by paired-end whole exome sequencing performed on Illumina HiSeq 2500 to achieve a minimal tumor coverage of 150X. Somatic variants were called using muTect (v1.1.7). The full pipeline is available here https://github.com/soccin/BIC-variants_pipeline. LMCB3 line was sequenced on Illumina Miseq using a targeted Truseq custom amplicon panel covering hotspot regions in pancreatic cancer driver genes: *KRAS, BRAF, CTNNB1, GNAS, CDKN2A, ARID1A, ARID1B, ARID2, KMT2A, KMT2B, KMT2C, ATM, SMAD3, SF3B1, RNF43, PCDH15, TGFBR1, TGFBR2, TP53, SMAD4*. Variants were called

using Illumina VariantStudio software. Somatic alterations were evaluated for their oncogenic and/or biologic impact using OncoKB [2] and literature curation.

Electromobility Shift Assay (EMSA)

Samples were prepared for EMSA using a canonical E-box motif as previously described [3] and DNA was detected using a 2nd generation DIG Gel Shift Kit (Roche).

Data Analysis

General. GNU parallel [4] was used where possible to increase efficiency of analysis.

RNA-seq. Reads were quality checked using FastQC v0.11.5 [5] and mapped to the human/mouse genome (hg19/mm10) with STAR2.5.2b [6] using standard settings for paired reads. On average, 85% of reads were uniquely mapped. Uniquely mapped reads were assigned to annotated genes with HTSeq v0.6.1p1 with default settings. Read counts were normalized by library size, and differential gene expression analysis based on a negative binomial distribution was performed using DESeq2 v3.4 [7]. Unless otherwise indicated, thresholds for differential expression were set as follows: adjusted p-value<0.05, fold change > 2.0 or < 0.5, and average normalized read count > 10. Gene set enrichment analysis was performed using GSVA v3.4 [8] or GSEA and previously curated gene sets [9].

ChIP-seq and ATAC-seq. Reads were quality checked using FastQC v0.11.5 and mapped to the human/mouse genome (hg19/mm10) with Bowtie2. Sam Tools was used to manipulate .sam and .bam files [10]. Tag directories, visualization in UCSC genome browser, and downstream analyses were performed using the HOMER suite [11].

Screens. Analysis of CRISPR screens was performed as previously described [12]. For the whole genome screen, after count table creation and filtering for expressed genes, guides were collapsed and genes tested for enrichment using CAMERA [13].

Analysis of public datasets. Count data were downloaded from the ICGC, GTEx, and Broad GDAC Firehose portals, respectively. For the cross-study comparisons, transcription factors were curated and ranked as previously described [14] and expression values were simulated based on rank-expression correlation. The top 5 factors in each case were selected where indicated, representing ~30% of the reads in each case. Principal components were calculated and plotted. Briefly, simulated expression levels were log-transformed and centered on the

mean, and eigenvectors were calculated based on the correlation matrix. All data manipulation and analyses were performed in R using base R packages unless otherwise specified. Code for the described analyses are available upon request.

ChIP-seq, RNA-seq, and microarray were downloaded from UCSC Genome Browser, or GEO when available [15-19]. In all other cases, fastq files were generated from the Sequence Read Archive using fastq-dump and re-analyzed using our in-house ChIPseq pipeline.

MSK-IMPACT [20], TCGA [21], and ICGC [22] oncogenomic data were accessed via cBioPortal [23, 24]. To enrich for possible driver mutations and reduce confounders introduced by high levels of passenger mutations, only pancreatic adenocarcinoma cases where <10% of the assayed genes were mutated and <50% contained copy number mutations (representing 95% of the cases) were included in the mutual exclusivity analysis.

Multispectral imaging, spectral unmixing and cell segmentation. Four color multiplex stained slides were imaged using the Vectra Multispectral Imaging System version 3 (Perkin Elmer). Scanning was performed at 20X (200X final magnification). Filter cubes used for multispectral imaging were DAPI, FITC, Cy3, Texas Red and Cy5. A spectral library containing the emitted spectral peaks of the fluorophores in this study was created using the Vectra image analysis software (Perkin Elmer). Using multispectral images from single-stained slides for each marker, the library was used to separate each multispectral cube into individual component (spectral unmixing) allowing for identification of all four marker channels of interest using Inform 2.4 image analysis software. Images were exported to Indica Labs Halo image analysis platform and cell segmentation and signal thresholding was performed separately on each case using a supervised algorithm.

Statistical tests. Statistical tests were performed in Prism 7 or R as appropriate. P values are reported based on two-sided, unpaired t tests, $\alpha=0.05$, unless otherwise specified. Bar graphs represent mean ± standard deviation unless otherwise specified.

Antibodies	**SOURCE**	**IDENTIFIER**
Rabbit mAb Anti-Phospho-Smad2 (S465/467)(138D4)	Cell Signaling Technology	Cat #3108
Rabbit mAb Anti-Smad2/3 (D7G7)	Cell Signaling Technology	Cat #8585
Rabbit mAb Anti-Mouse/Human Id1 (BCH-1/195-14)	Biocheck	Cat #BCH1/195-14
Rabbit mAb Anti-Mouse Id1 (BCH-1/37-2)	Biocheck	Cat #BCH-1/37-2
Rabbit mAb Anti-E-Cadherin (24E10)	Cell Signaling Technology	Cat #3195
Rabbit pAb Anti-Klf5 (ab137676)	Abcam	Cat #ab137676

Rabbit mAb Anti-Smad4 (ab40759)	Abcam	Cat #ab40759
Rabbit mAb Anti- Phospho-Akt (Thr308) (D25E6)	Cell Signaling Technology	Cat #13038
Rabbit mAb Anti- Phospho-Akt (Ser473) (D9E)	Cell Signaling Technology	Cat #4060
Mouse mAb Anti-Akt (pan) (40D4) Mouse	Cell Signaling Technology	Cat #2920
Rabbit mAb Anti- Phospho-Smad1/5 (S463/465) (41D10)	Cell Signaling Technology	Cat #9516
Rabbit mAb Anti-Vimentin (D21H3)	Cell Signaling Technology	Cat #5741
Rabbit mAb Anti-FOXO1	Cell Signaling Technology	Cat #2880
Rabbit pAb Anti-FOXO3a	Abcam	Cat #ab12162
Rabbit mAb Anti-ID1	Biocheck	Cat#195-14
Rabbit pAb Anti-SOX4	Diagenode	Cat# C15310129
Rabbit pAb Anti-goat IgG	Vector lab	Cat#PI-9500
Rabbit pAb Anti-E2A	Santa Cruz	Cat#SC-349
Mouse mAb Anti-SOX4 (Clone 5634)	Abcam	Cat#Ab243041
Mouse mAb Anti-PanCK (Clone Ae1/Ae3)	Dako	Cat#M351501-2
Mouse mAb Anti-Cytokeratin (Clone Cam5.2)	BD Biosciences	Cat#349205
Mouse mAb Anti-N-Cadherin (13A9)	Cell Signaling Technology	Cat #14215
Mouse mAb Anti-CK7 (Clone OV-TL-12/30)	Invitrogen	Cat#MA5-11986
Mouse mAb Anti-α-Tubulin (DM1A)	Sigma	Cat #T6199
Mouse mAb Anti-RNA polymerase II CTD repeat YSPTSPS (8WG16)	Abcam	Cat #ab817
Rat mAb Anti-CXCR4-PE	Biolegend	Cat #146505
Mouse mAb Anti-SSEA4-FITC	Biolegend	Cat #330409
Rat mAb Anti-CD44-PE	eBioscience	Cat #12-0441-83
Rat mAb Anti-CD24-FITC	eBioscience	Cat #11-0241-81
Rat mAb Anti-CD45-BUV395	BD Biosciences	Cat #565967
Rat mAb Anti-mouse CD16/32	Biolegend	Cat #146505
Rat mAb Anti-CK19	DSHB	Cat #TROMAIII
Goat pAb Anti-KLF5	R & D Systems	Cat#AF3758

Oligonucleotides to generate sgRNAs for Crispr-Cas9 mediated knockout

Name	Top_oligo	Bottom_oligo
*Id1*_c-term	caccgGATCGCATCTTGTGT CGCTG	aaacCAGCGACACAAGATG CGATCc
NonTargeting_0001_MGLibA_66406	CACCgcgaggtattcggctccgcg	AAACcgcggagccgaatacctcgc
NonTargeting_0002_MGLibA_66407	CACCgctttcacggaggttcgacg	AAACcgtcgaacctccgtgaaagc
NonTargeting_0003_MGLibA_66408	CACCgatgttgcagttcggctcgat	AAACatcgagccgaactgcaacatC
Region2__3	CACCgaacgttctgaacccgccct	AAACagggcgggttcagaacgttc
*Tgfbr2*_MGLibA_53621	CACCgaccgcaccgccattgtcgc	AAACgcgacaatggcggtgcggtc
*Tgfbr2*_MGLibA_53622	CACCgccttgtagacctcggcgaag	AAACcttcgccgaggtctacaaggC
*Tgfbr2*_MGLibA_53623	CACCgccacgcgaagggcaacctgc	AAACgcaggttgcccttcgcgtggC

Oligonucleotides to generate shRNAs

Name	shRNA_oligo
mId1.840	TGCTGTTGACAGTGAGCGATGAAAATATTGTTTTACAATATAGTGAAGCCACAGATGTA TATTGTAAAACAATATTTTCAGTGCCTACTGCCTCGGA
mId3.812	TGCTGTTGACAGTGAGCGACCCTGATTATGAACTCTATAATAGTGAAGCCACAGATGT ATTATAGAGTTCATAATCAGGGCTGCCTACTGCCTCGGA
mId2.91	TGCTGTTGACAGTGAGCGCCCCGATGAGTCTGCTCTACAATAGTGAAGCCACAGATGT ATTGTAGAGCAGACTCATCGGGTTGCCTACTGCCTCGGA
Renilla.713	TGCTGTTGACAGTGAGCGCAGGAATTATAATGCTTATCTATAGTGAAGCCACAGATGT ATAGATAAGCATTATAATTCCTATGCCTACTGCCTCGGACTTCAAGGGGCTAGAATTCC TGTAGTAGGTGGCTGAGGC

qPCR Primers

Gene	Forward Primer	Reverse Primer
hSMAD7	TTCCTCCGCTGAAACAGGG	CCTCCCAGTATGCCACCAC
hGAPDH	CTGGGCTACACTGAGCACC	AAGTGGTCGTTGAGGGCAATG
hID1	CTGCTCTACGACATGAACGG	GAAGGTCCCTGATGTAGTCGAT
hAHR	CAAATCCTTCCAAGCGGCATA	CGCTGAGCCTAAGAACTGAAAG
hKLF5	CCTGGTCCAGACAAGATGTGA	GAACTGGTCTACGACTGAGGC
hHMGB2	CCGGACTCTTCCGTCAATTTC	GTCATAGCGAGCTTTGTCACT
hPPARG	ACCAAAGTGCAATCAAAGTGGA	ATGAGGGAGTTGGAAGGCTCT
mAhr	GCCCTTCCCGCAAGATGTTAT	TCAGCAGGGGTGGACTTTAAT
mKlf5	CAGGCCACCTACTTTCCCC	GAATCGCCAGTTTGGAAGCAA
mHmgb2	CGGGGCAAAATGTCCTCGTA	ATGGTCTTCCATCTCTCGGAG
mPparg	GGAAGACCACTCGCATTCCTT	GTAATCAGCAACCATTGGGTCA
mId1_1	CCTAGCTGTTCGCTGAAGGC	CTCCGACAGACCAAGTACCAC
mId1_4	GCGAGGTGGTACTTGGTCTG	AGGATCTCCACCTTGCTCAC
mSmad7	GGCCGGATCTCAGGCATTC	TTGGGTATCTGGAGTAAGGAGG
mGapdh	AGGTCGGTGTGAACGGATTTG	TGTAGACCATGTAGTTGAGGTCA
mSnai1	CACACGCTGCCTTGTGTCT	GGTCAGCAAAAGCACGGTT
Id1_peak1	AGAATGCTCCAGCCCAGTTT	GAGATCACCCCCTCCCTTT
Id1_peak2	CTCCGCCTGTTCTCAGGAT	CGTAGAGCAGGACGTTCACC
mRnf152_enhancer	CGCTTACTGAGCACCACAAA	TGTCGTCCAGGACAGTCTGA
mFbxo32_enhancer	CACAGGTGGCCAGTGTATTG	TGTCCTTTTGTCCCTTCCTG

Primers for cloning

*Id1*_HAL	ATATATgaattcAATAAGTAGAGATCACAGCC	ATATATggtaccGCGACACAAGATGCGATCGTC
*Id1*_HAR	ATATATggatccGGCGGCGCACTGAGGGACCAG	ATATATtctagaCTTTAGAACCCATCTGTGCC
eGFP	ATATATggtaccGTGAGCAAGGGCGAGGAGCTGTTCA	ATATATggatccCTACTTGTACAGCTCGTCCATGCCGAGA
Id1	GATAGggatccATGAGGTCGCCAGTGGCAGTG	AAGCTgaattcTCAGCGACACAAGATGCGATCGTC

Citations

1. Ran, F.A., et al., *Genome engineering using the CRISPR-Cas9 system.* Nat Protoc, 2013. **8**(11): p. 2281-2308.

2. Chakravarty, D., et al., *OncoKB: A Precision Oncology Knowledge Base.* JCO Precis Oncol, 2017. **2017**.

3. Kreider, B.L., et al., *Inhibition of myeloid differentiation by the helix-loop-helix protein Id.* Science, 1992. **255**(5052): p. 1700-2.

4. Tange, O., *GNU Parallel - The Command-Line Power Tool*. 2011: The USENIX Magazine.

5. Andrews, S., *FastQC*, in *Babraham Bioinformatics*. 2010.

6. Dobin, A., et al., *STAR: ultrafast universal RNA-seq aligner.* Bioinformatics, 2013. **29**(1): p. 15-21.

7. Love, M.I., W. Huber, and S. Anders, *Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2.* Genome Biol, 2014. **15**(12): p. 550.

8. Hanzelmann, S., R. Castelo, and J. Guinney, *GSVA: gene set variation analysis for microarray and RNA-seq data.* BMC Bioinformatics, 2013. **14**: p. 7.

9. Subramanian, A., et al., *Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles.* Proc Natl Acad Sci U S A, 2005. **102**(43): p. 15545-50.

10. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. **25**(16): p. 2078-9.

11. Heinz, S., et al., *Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities.* Mol Cell, 2010. **38**(4): p. 576-89.

12. Sanjana, N.E., O. Shalem, and F. Zhang, *Improved vectors and genome-wide libraries for CRISPR screening.* Nat Methods, 2014. **11**(8): p. 783-784.

13. Wu, D. and G.K. Smyth, *Camera: a competitive gene set test accounting for inter-gene correlation.* Nucleic Acids Res, 2012. **40**(17): p. e133.

14. David, C.J., et al., *TGF-beta Tumor Suppression through a Lethal EMT.* Cell, 2016. **164**(5): p. 1015-30.

15. Moffitt, R.A., et al., *Virtual microdissection identifies distinct tumor- and stroma-specific subtypes of pancreatic ductal adenocarcinoma.* Nat Genet, 2015. **47**(10): p. 1168-78.

16. Bernstein, B.E., et al., *The NIH Roadmap Epigenomics Mapping Consortium.* Nat Biotechnol, 2010. **28**(10): p. 1045-8.

17. Diaferia, G.R., et al., *Dissection of transcriptional and cis-regulatory control of differentiation in human pancreatic cancer.* EMBO J, 2016. **35**(6): p. 595-617.

18. Kent, W.J., et al., *The human genome browser at UCSC.* Genome Res, 2002. **12**(6): p. 996-1006.

19. Rosenbloom, K.R., et al., *ENCODE data in the UCSC Genome Browser: year 5 update.* Nucleic Acids Res, 2013. **41**(Database issue): p. D56-63.

20. Zehir, A., et al., *Mutational landscape of metastatic cancer revealed from prospective clinical sequencing of 10,000 patients.* Nat Med, 2017. **23**(6): p. 703-713.

21. TCGA, *Integrated Genomic Characterization of Pancreatic Ductal Adenocarcinoma.* Cancer Cell, 2017. **32**(2): p. 185-203 e13.

22. Biankin, A.V., et al., *Pancreatic cancer genomes reveal aberrations in axon guidance pathway genes.* Nature, 2012. **491**(7424): p. 399-405.

23. Cerami, E., et al., *The cBio cancer genomics portal: an open platform for exploring multidimensional cancer genomics data.* Cancer Discov, 2012. **2**(5): p. 401-4.

24. Gao, J., et al., *Integrative analysis of complex cancer genomics and clinical profiles using the cBioPortal.* Sci Signal, 2013. **6**(269): p. pl1.