www.sciencemag.org/cgi/content/full/science.1200801/DC1

Supporting Online Material for

Aberrant Overexpression of Satellite Repeats in Pancreatic and Other Epithelial Cancers

David T. Ting, Doron Lipson, Suchismita Paul, Brian W. Brannigan,Sara Akhavanfard, Erik J. Coffman, Gianmarco Contino, Vikram Deshpande, A. John Iafrate, Stan Letovsky, Miguel N. Rivera, Nabeel Bardeesy, Shyamala Maheswaran, Daniel A. Haber*

*To whom correspondence should be addressed. E-mail: haber@helix.mgh.harvard.edu

Published 13 January 2011 on *Science* Express
DOI: 10.1126/science.1200801

This PDF file includes:

MATERIALS AND METHODS

Tissues and Cell lines

Mice with pancreatic cancer of different genotypes were bred as previously described in the Bardeesy laboratory (*S1*). Normal wild type mice were purchased from Jackson laboratories. Animals were euthanized as per animal protocol guidelines. Pancreatic tumors and normal tissue were extracted sterilely and then flash frozen with liquid nitrogen. Tissues were stored at -80° C. Additional normal total RNA was purchased from Clontech (Mouse Total RNA Master Panel). Cell lines were generated fresh for animals as previously described (*S2*) and established cell lines were cultured in RPMI-1640 + 10% FBS + 1% Pen/Strep (Gibco/Invitrogen). Cell lines treated with 5-azacytidine (AZA) was done at a concentration of 10 μM for 48 hrs before RNA extraction. Additional mouse tumors from colon and lung were generously provided by Kevin M. Haigis (Massachusetts General Hospital) and Kwok-Kin Wong (Dana-Farber Cancer Institute).

Human pancreatic tumor tissues were obtained by V. Deshpande as excess discarded human material per IRB protocol from the Massachusetts General Hospital. Additional pancreatic, lung, renal, ovarian, and prostate tumors were obtained from the Massachusetts General Hospital Tissue Bank Repository. Gross tumor was excised and fresh frozen in liquid nitrogen prior to nucleic acid extraction. Two independent normal human pancreatic tissue total RNAs were obtained from Clontech and Ambion. Other normal human tissue RNA was obtained from commercial vendors, Clontech (Human Total RNA Master Panel II).

Nucleic Acid Preparation

Fresh frozen tissue was pulverized with a sterile pestle in a microfuge tube on dry ice. Cell lines were cultured and fresh frozen in liquid nitrogen prior to nucleic acid extraction. RNA and DNA from cell lines and fresh frozen tumor and normal tissues were all processed in the same manner. RNA was extracted using the TRIzol® Reagent (Invitrogen) per manufacturer's specifications. DNA from tissue and cell lines was extracted using the QIAamp Mini Kit (QIAGEN) per manufacturer's protocol.

Single Molecule Sequencing and Data Processing

Purified RNA was subjected to Digital Gene Expression (DGE) sample prepping and analysis on the HeliScope™ Single Molecule Sequencer from Helicos BioSciences. This method has been previously described (*S3*). Briefly, Single stranded cDNA was reverse transcribed from RNA with a dTU25V primer and the Superscript III cDNA synthesis kit (Invitrogen). RNA was digested and single stranded cDNA was purified using a solid phase reversible immobilization (SPRI) technique with Agencourt® AMPure® magnetic beads. Single stranded cDNA was denatured and then a poly-A tail was added to the 3' end using terminal transferase (New England Biolabs).

Purified DNA was subjected to the DNA sequencing sample prepping protocol from Helicos that has been previously described (*S4*). Briefly, genomic DNA was sheared with a Covaris S2 acoustic sonicator producing fragments averaging 200 bps and ranging from 100-500 bps. Sheared DNA was then cleaned

with SPRI. DNA was then denatured and a poly-A tail was added to the 3' end using terminal transferase.

Tailed cDNA or DNA were then hybridized to the sequencing flow cell followed by "Fill and Lock" and single molecule sequencing. Sequence reads were subjected to filtering for a minimum read length of 25 and removal of artifact read followed by alignment to the known human or mouse transcriptome libraries (UCSC knownGene database +rRNA sequences) using the indexDPgenomic aligner (*S3*). Genomic DNA sequence reads were aligned to the human (UCSC hg18) or mouse (UCSC mm9) genomes and counted to determine copy number of the major mouse satellite.

Northern Blot Analysis

Northern blot probe for major mouse satellite was generated by performing PCR of genomic DNA around a 788 bp tandem array insertion of the mouse major satellite at chromosome 2: 98506702:98507489 (Build: UCSC mm9). Primers used had the following sequences:

Forward: CGTTTCCAACGAATGTGTTT at chr2: 98506541-9850650
Reverse: TGGAAACAGATGATTTCGTC at chr2: 98507480-98507499

The resulting PCR product was approximately 958 bp. The PCR product was then cloned with the TOPO® TA Cloning Dual Promoter Kit (Invitrogen) per protocol.

Northern blot was performed using the NorthernMax-Gly Kit (Ambion). Total RNA (10 ug) was mixed with equal volume of Glyoxal Load Dye (Ambion) and incubated at 50°C for 30 min. After electrophoresis in a 1% agarose gel, RNA was transferred onto BrightStar-Plus membranes (Ambion) and crosslinked with ultraviolet light. The membrane was prehybridized in ULTRAhyb buffer (Ambion) at 68°C for 30 min. The mouse RNA probe (1100 bp) was generated from the pCRII-TOPO vector in both sense (T7) and anti-sense directions (SP6) using the MAXIscript Kit (Ambion). The RNA probe was then nonisotopically labeled using the BrightStar Psoralen-Biotin Kit (Ambion) according to the manufacturer's instructions. Using 0.1 nM probe, the membrane was hybridized in ULTRAhyb buffer (Ambion) at 68°C for 2 hours. The membrane was washed with a Low Stringency wash at room temperature for 10 min, followed by two High Stringency washes at 68°C for 15 min. For nonisotopic chemiluminescent detection, the BrightStar BioDetect Kit was used according to the manufacturer's instructions.

RNA *in situ* hybridization

Formalin Fixed Paraffin Embedded (FFPE) tissue from human and mouse was prepared per standard protocol by the Massachusetts General Hospital Clinical and Research Pathology Cores. Human FFPE tissues were selected by V. Deshpande according to IRB protocols. Pancreatic tissue FFPE cases were all obtained retrospectively. We note that samples were not collected with the intent for RNA-ISH and that there is variability in nucleic acid preservation.

For mouse tissues, ISH was performed using standard protocols. Briefly, dissected tissues were fixed in 4% paraformaldehyde and embedded in paraffin blocks. 8 micron sections were deparaffinized and digested with pepsin (Digest-All 3, Invitrogen), treated with acetic anhydride and dehydrated with ethanol. Sections were hybridized overnight with sense and anti-sense digoxigenin labeled probes prepared with the DIG RNA Labeling Kit (Roche). Since transcript was found in both sense and antisense directions, a probe against WT1 was used as a negative control. Bound probes were detected using an

alkaline phosphatase conjugated anti-digoxigenin antibody (Roche) and BM purple (Roche). Counter stain applied was nuclear fast red (Vector laboratories). The probe used is the same used for northern blot analysis as above. Images were taken with an Olympus BX60 scope with color camera.

For human tissues, ISH was performed using Panomics protocols. 5 micron sections were cut, fixed in 10% formaldehyde (Fisher Scientific, Pittsburgh, PA), deparaffinized, boiled in pre-treatment solution (Affymetrix, Santa Clara, CA) and digested with proteinase K (Affymetrix, Santa Clara, CA). Sections were hybridized for 3hour at 40° with a custom designed Panomics QuantiGene ViewRNA probe against the HSATII satellite (Affymetrix, Santa Clara, CA). Bound probes were then amplified per protocol from Panomics using PreAmp and Amp molecules. Multiple Label Probe oligonucleotides conjugated to alkaline phosphatase (LP-AP) are then added and Fast Red Substrate is used to produce signal (red dots). Slides are then counterstained with Hematoxylin. Serial sections were also subjected to Hematoxylin & Eosin staining per standard histology protocol. Images were taken by a Nikon 90i scope with color camera.

Gene Expression and Satellite Expression Profiling

All mouse sample reads were aligned to a custom made library for the mouse major satellite (Sequence from UCSC genome browser). Human samples were aligned to a custom made reference library for all satellite repeats and LINE-1 variants generated from the Repbase library (*S5*). Reads aligning to the satellite repeat reference were counted and their fraction of the total number of genome aligned reads was determined. In addition, gene expression counts were determined based on alignments to the annotated human/mouse transcriptomes based on a statistical gene counting method (*S*3). Gene counts were normalized per million aligned genomic reads for all samples.

For linear correlation of mouse major satellite and human alpha (ALR) satellite to transcriptome we performed a linear regression analysis of the normalized expression levels of each gene across all tissues vs the measured satellite expression level. Genes were then ranked according to decreasing Pearson coefficient for linear regression, and a set of highly-correlating genes ($R > 0.85$) was selected for further analysis.

Gene Categorization

Linearly correlated genes with $R > 0.85$ were mapped using the DAVID program (*S6, S7*). These genes were then analyzed with the Functional Annotation clustering program and the UP_TISSUE database to classify each of these mapped genes. Germ/Stem cell genes included genes expressed highly in testis, egg, trophoblast, and neural stem cells. Neural genes included genes expressed highly in brain, spinal cord, and specialized sensory neurons including olfactory, auditory, and visual perception. Zinc Finger proteins were classified using the INTERPRO database.

Genomic proximity of LINE-1

The transcriptional start sites for all mouse genes were determined (UCSC knownGene database) as well as the position of all LINE-1 elements in the mouse and human genomes with a minimum of 1 Kbp in length (UCSC RepeatMasker database). For all genes with a minimum average expression level of 1 transcripts per million, the distance to the closest upstream LINE-1 element was calculated. Focusing on the top satellite correlated genes (SCG; $R > 0.85$), we calculated the frequency of these genes from the

total gene set as a function of the distance from the closest upstream LINE-1 element. Enrichment factor and p-value according to Fisher's Exact test were calculated at a distance of 10 Kbp.

Immunohistochemistry for Chromogranin A

Immunohistochemistry for mouse tumor and normal tissue sections was performed using standard protocols with an anti-Chromogranin A antibody (Abcam ab45138, 1:200). Briefly, sections were de-paraffinized using Xylene for 30 minutes and antigen retrieval was performed using Borg Decloaking solution in a Decloaking Chamber (Biocare Medical). Tissues were incubated with primary antibodies for 2 hours and with secondary antibodies for 1 hour at room temperature. Signals were detected using the ABC kit for immunoperoxidase staining (Vector laboratories). Images were taken by a Nikon 90i scope with color camera.

Human neuroendocrine marker targeted differential expression analysis

A targeted gene expression analysis of known neuroendocrine was carried out in human PDAC samples. Human tumors were separated into high vs low ALR satellite levels using the median satellite expression. A total of 12 genes were evaluated between high and low satellite tumors and fold change was calculated. Analysis of the population means was compared using the 2-tailed student t-test assuming unequal variance and p-values were corrected to q-values using the Storey method (*S8*). Genes that had q-values < 0.05 were considered significant candidates.

Histone modifiers targeted differential expression analysis

A targeted gene expression analysis of demethylases, methyltransferases, acetyltransferases, and deacetylases was carried out in mouse and human PDAC samples. A list of these modifiers was obtained from three recent (*S9-S11*). Mouse PDACs with $Kras^{G12D}$ and *Tp53* loss and all human PDACs were used for this analysis. Mouse and human tumors were separated into high vs low satellite levels using the median satellite expression (Mouse major: 56,081 tpm and Human ALR: 13,428 tpm). A total of 99 genes were evaluated between high and low satellite tumors and fold change was calculated. Analysis of the population means was compared using the 2-tailed student t-test assuming unequal variance and p-values were corrected to q-values using the Storey method (*S8*). Genes that had q-values < 0.10 were considered significant candidates.

SUPPLEMENTARY TEXT

Presence of specific mouse and human satellite sequences in parasitic genomes.

BLAST sequence matching of overexpressed mouse tumor satellites first identified multiple independent sequences (max identity of >95% , 100% coverage, E-score < $2x10^{-176}$) in three murine malaria strains (*Plasmodium berghei* (461 hits), *Plasmodium yoelii yoelii* (69 hits), and *Plasmodium chabaudi chabaudi* (44 hits)), as well as the parasitic trematode *Schistosoma mansoni* (20 hits). Similarly, BLAST analysis demonstrated presence of the human satellite HSAT6 (72-85% max identity, 100% coverage, E-value of $2x10^{-5}$) in human malaria *Plasmodium falciparum*, but not in other malarial species. Interestingly, the second hit for *Plasmodium falciparum* is identified by NCBI as a conserved Plasmodium protein without known function that maps to *P. falciparum* chromosome 12. It is possible that these unexpected sequence alignments result from contamination of parasitic DNA with the abundant satellites from the host used to cultivate these species *in vitro*, although HSAT6 is not the most prevalent human satellite DNA, and other more abundant subtypes were not identified. The possibility of horizontal transfer of host-specific satellite sequences to these parasitic strains needs to be excluded, particularly given the unique role of ncRNAs in the regulation of gene expression in Plasmodium species (*S12*)

SUPPLEMENTARY FIGURES

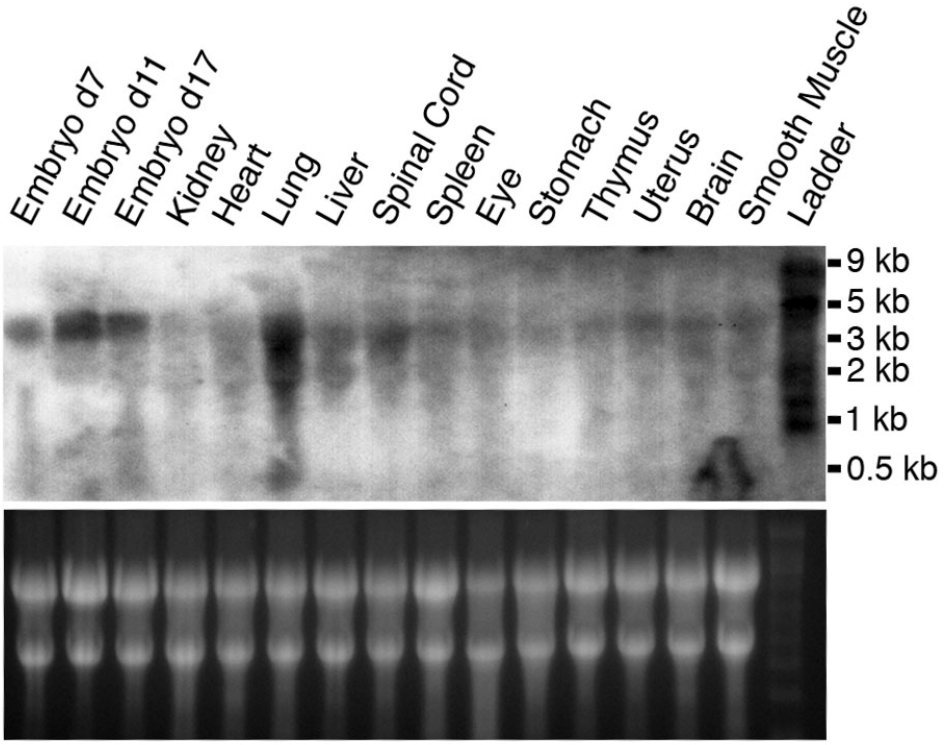

Figure S1: Northern blot analysis from normal mouse tissues
Total RNA isolated from multiple adult and fetal mouse tissues analyzed by northern blot.

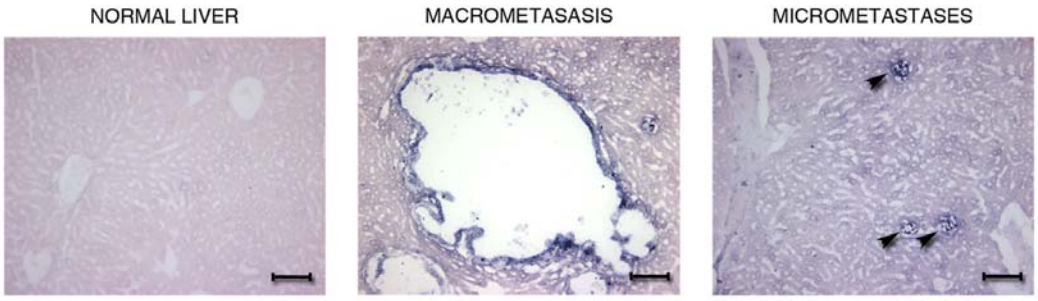

Figure S2: RNA-ISH of metastatic mouse PDAC
PDAC metastatic lesions in liver, which itself does not express satellites (left). Large, glandular metastatic tumor deposits are readily identified by standard histological evaluation and stain for satellite (middle). Satellite expressing micrometastases in liver (right; arrowheads). 200x magnification (scale bar = 100 μm).

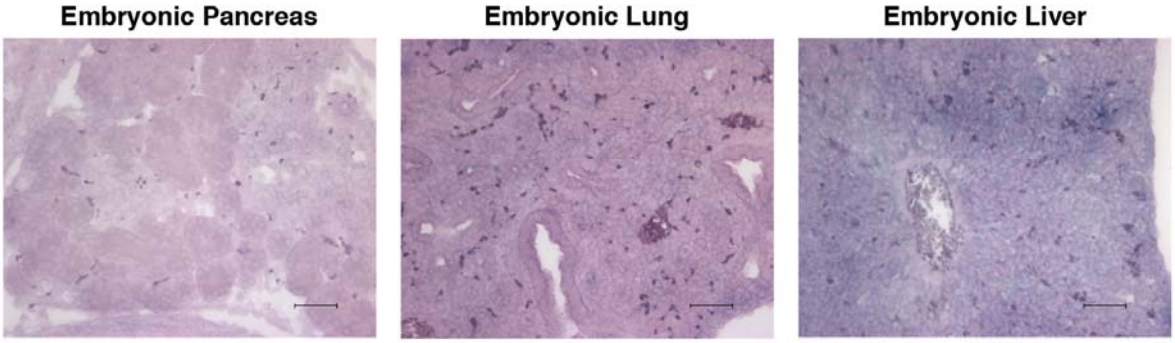

Figure S3: RNA-ISH of embryonic organs
Review of whole embryo day 15 demonstrated strongest signal in embryonic liver and lung. Images of embryonic pancreas, liver, and lung shown at 400x magnification (scale bar = 50 μm).

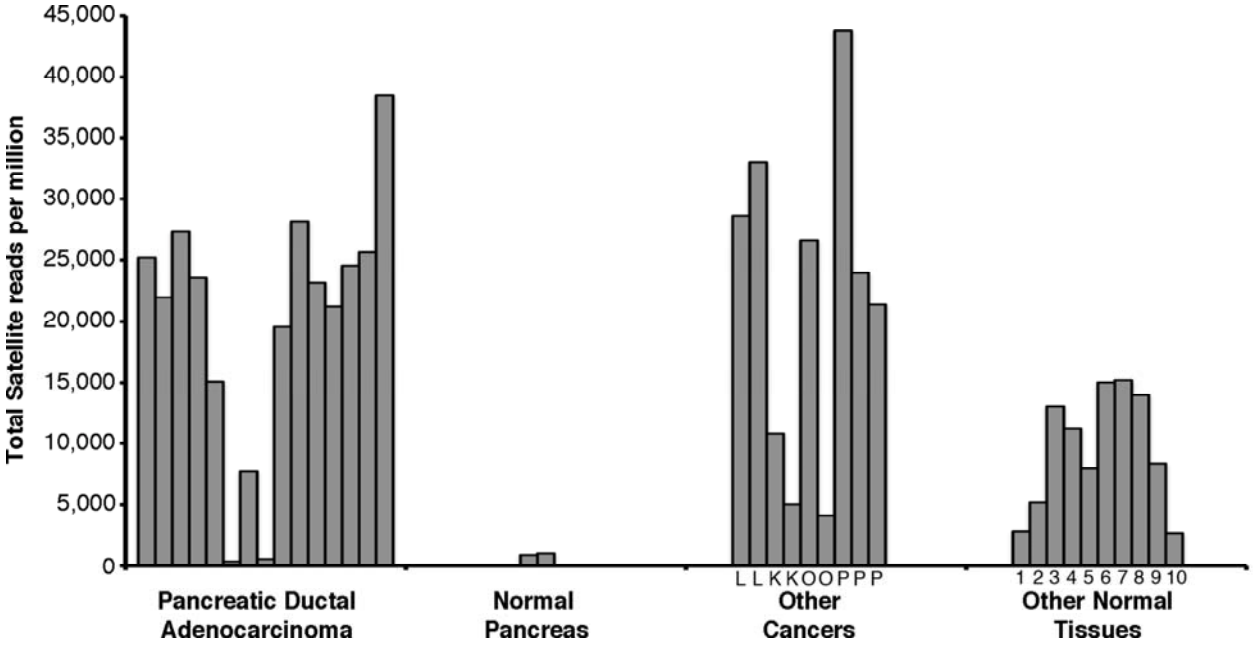

Figure S4: Total satellite expression in human tissue
Human pancreatic ductal adenocarcinoma, normal pancreas, other cancers (L – lung, K – kidney, O – ovary, P – prostate), and other normal human tissues (1 - fetal brain, 2 – adult brain, 3 - colon, 4 - fetal liver, 5 – adult liver, 6 - lung, 7 - kidney, 8 - placenta, 9 - prostate, and 10 - uterus) quantitated by DGE. Satellite expression is shown as transcripts per million aligned to human genome.

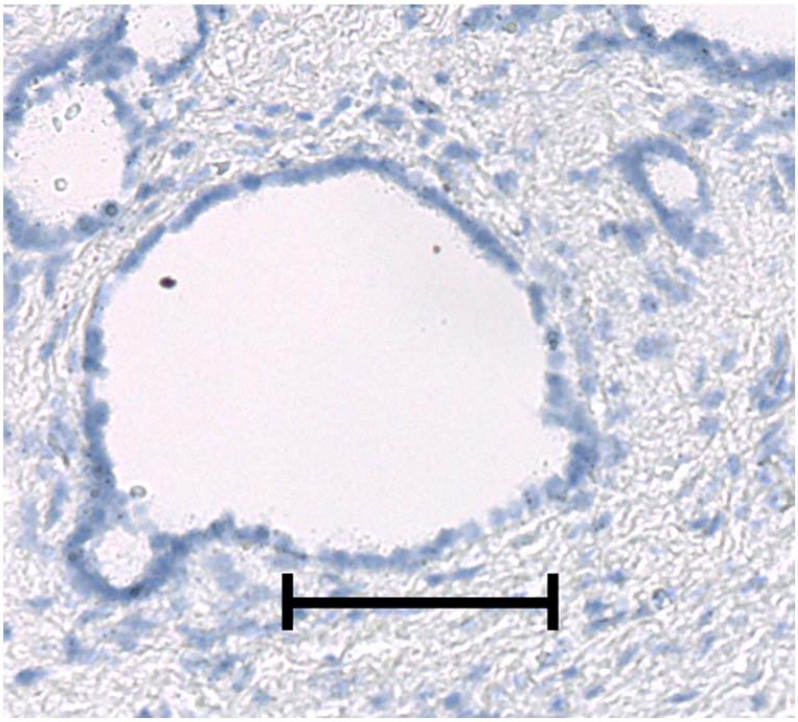

Figure S5: HSATII RNA-ISH in chronic pancreatitis
A total of 8 FFPE blocks of chronic pancreatitis were subjected to RNA-ISH for HSATII with nearly absent signal. This absence of signal is different than the normal adjacent tissues we observed in PDAC blocks and may reflect true differences in biology or differences in nucleic acid integrity in chronic pancreatitis.

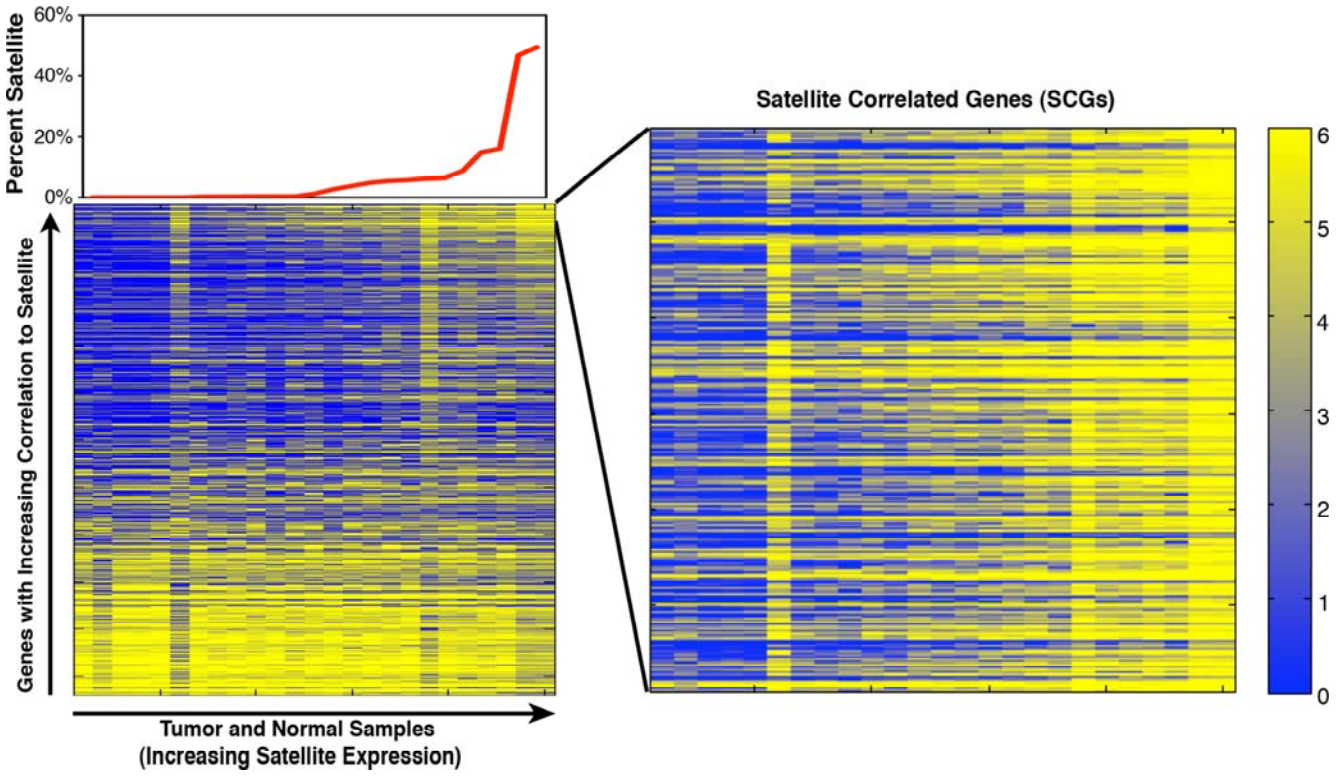

Figure S6: Linear correlation of transcriptome to major mouse satellite
Multiple linear correlation analysis of mouse major satellite to other cellular transcripts in all mouse tumors and normal tissues, depicted as heat map (X-axis, samples ordered by increasing expression of major satellite; Y-axis, genes ordered by linear correlation to major satellite expression), with yellow (high) and blue (low) shown as log2 (reads per million). Expanded view of top satellite correlated genes (SCGs) with highest linearity ($R > 0.85$) with satellite levels.

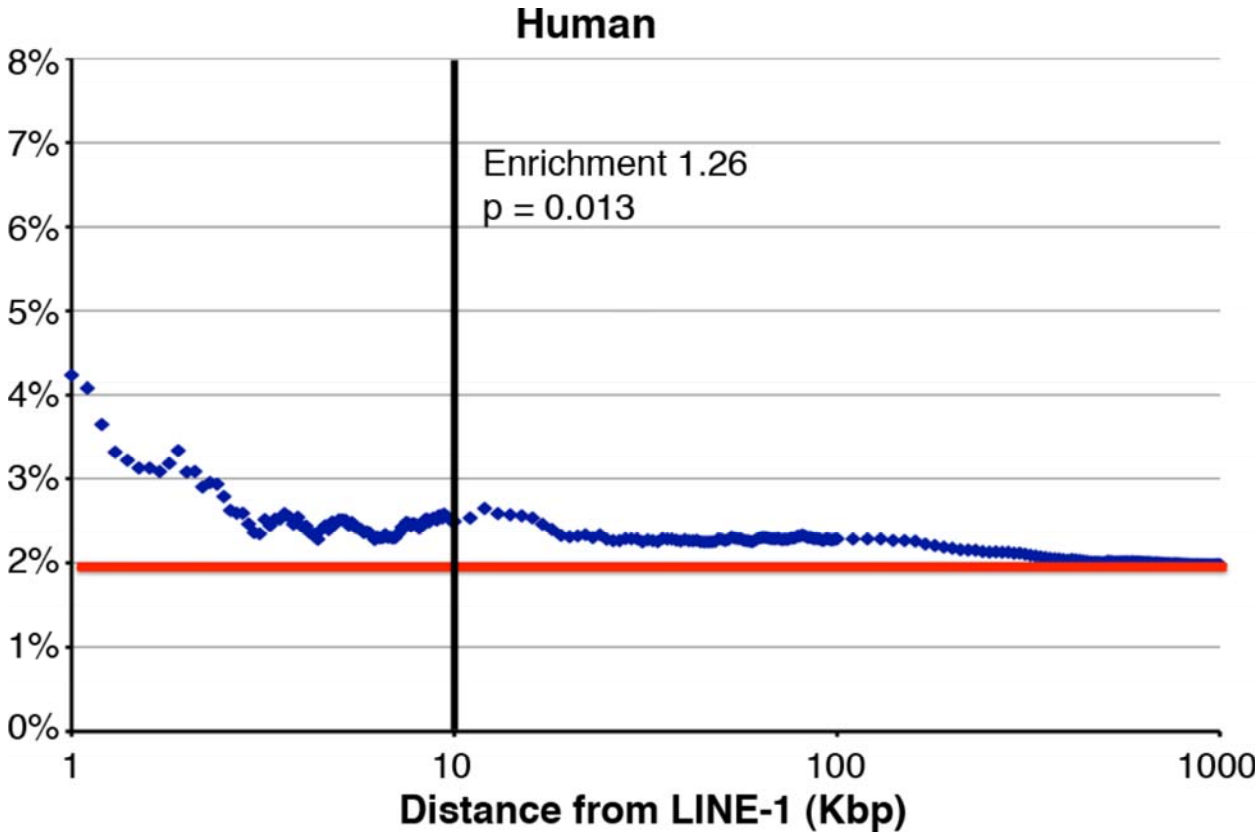

Figure S7: Relationship of human SCGs to distance from LINE-1
Fraction of SCGs (blue) versus predicted (red) transcriptional start sites within a given distance of a LINE-1 insertion. Enrichment calculations were done at a distance of 10 kbp (black line)

SUPPLEMENTARY TABLES

Table S1: Mouse tissue digital gene expression
Mouse samples subjected to sequencing with total genomic reads and percentage of reads aligning to transcriptome and major satellite among multiple mouse tumors, cell lines, and normal tissues.

Mouse ID	Tissue Type	Genotype	Total Genomic Reads	% Transcriptome Reads	% Major Satellite Reads
AH284 Rep 1	Pancreatic Cancer	$Kras^{G12D}$, $Tp53^{lox/+}$	18,063,363	9.56%	46.84%
AH284 Rep 2	Pancreatic Cancer	$Kras^{G12D}$, $Tp53^{lox/+}$	16,948,693	10.15%	49.54%
AH284 – 2*	Pancreatic Cancer	$Kras^{G12D}$, $Tp53^{lox/+}$	1,613,592	48.67%	4.78%
AH287	Pancreatic Cancer	$Kras^{G12D}$, $Tp53^{lox/+}$	2,227,850	54.70%	0.07%
AH288	Pancreatic Cancer	$Kras^{G12D}$, $Tp53^{lox/+}$	6,780,821	26.57%	14.79%
AH291	Pancreatic Cancer	$Kras^{G12D}$, $Tp53^{lox/+}$	1,388,906	43.12%	1.22%
AH294	Pancreatic Cancer	$Kras^{G12D}$, $Tp53^{lox/+}$	969,896	37.20%	3.73%
AH323	Pancreatic Cancer	$Kras^{G12D}$, $SMAD4^{lox/lox}$	1,887,663	72.73%	0.29%
AH346	Pancreatic Cancer	$Kras^{G12D}$, $SMAD4^{lox/lox}$	1,291,648	32.92%	6.07%
AH347	Pancreatic Cancer	$Kras^{G12D}$, $SMAD4^{lox/lox}$	1,634,314	38.94%	8.59%
AH348	Pancreatic Cancer	$Kras^{G12D}$, $SMAD4^{lox/lox}$	2,030,197	45.84%	5.61%
Colon 1	Colon Cancer - 1	$Kras^{G12D}$, $APC^{lox/lox}$	2,954,930	77.49%	0.07%
Colon 1	Colon Cancer - 2	$Kras^{G12D}$, $APC^{lox/lox}$	985,510	53.13%	6.27%
Colon 1	Colon Cancer - 3	$Kras^{G12D}$, $APC^{lox/lox}$	1,017,319	30.71%	16.02%
KN2128	Lung Cancer	$Kras^{G12D}$, $Tp53^{lox/lox}$	2,233,183	60.78%	2.66%
KN2199	Lung Cancer	$Kras^{G12D}$, $Tp53^{lox/lox}$	1,653,948	43.21%	5.37%
AH323	PDAC Cell Line	$Kras^{G12D}$, $SMAD4^{lox/lox}$	1,958,108	83.13%	0.02%
AH324	PDAC Cell Line	$Kras^{G12D}$, $Tp53^{lox/+}$	3,301,108	86.32%	0.04%
NB490	PDAC Cell Line	$Kras^{G12D}$, $Tp53^{lox/lox}$	15,378,802	76.85%	0.03%
AH284 Rep 1	Matched Normal Liver	$Kras^{G12D}$, $Tp53^{lox/+}$	2,270,669	75.68%	0.40%
AH284 Rep 2	Matched Normal Liver	$Kras^{G12D}$, $Tp53^{lox/+}$	1,627,749	56.59%	0.34%
AH284 – 2*	Matched Normal Liver	$Kras^{G12D}$, $Tp53^{lox/+}$	644,316	41.10%	0.31%
Colon 1	Matched Normal Liver	$Kras^{G12D}$, $APC^{lox/lox}$	1,536,346	86.53%	0.02%
Normal 1	Normal Pancreas	WT	247,582	14.49%	0.41%
Normal 2	Normal Pancreas	WT	244,719	11.15%	0.41%

* AH284-2 was RNA extraction from a different part of the pancreatic tumor and liver

Table S2: Genomic DNA copy number variation analysis of AH284 PDAC and matched liver.
To determine whether genomic amplification of satellite repeats also contributes toward the exceptional abundance of these transcripts in mouse pancreatic tumors, we analyzed the index AH284 tumor using next generation DNA digital copy number variation (CNV) analysis: satellite DNA comprise 18.8% of all genome-aligned reads in this tumor, compared with 2.3% of genomic sequences in matched normal liver. The major satellite repeat has previously been estimated at approximately 3% of the normal mouse genome (*13*). Thus, in this tumor with >100-fold increased expression of satellite repeats, approximately 8-fold gene amplification of the repeats may contribute to their abnormal expression. Major satellite reads and percent of total genomic reads shown.

	Major Satellite Reads	Total Genomic Reads
AH284 Liver	183,327 (2.3%)	7,995,538
AH284 PDAC	2,283,436 (18.8%)	12,124,201

Table S3: Human tissue digital gene expression

Human samples subjected to sequencing with total genomic reads and proportion of reads aligning to all human satellites, alpha (ALR) satellite, and HSATII satellite in transcripts per million (tpm).

SAMPLE ID	Genome	Total Satellite (tpm)	ALR (tpm)	HSATII (tpm)
PDAC 1	4,472,810	25,209	14,688	3,589
PDAC 2	1,668,281	22,001	12,653	3,295
PDAC 3	5,211,399	27,366	15,921	5,057
PDAC 4	1,649,041	23,556	13,428	3,167
PDAC 5	239,483	15,095	8,259	509
PDAC 6	1,520,470	374	195	14
PDAC 7	1,449,321	7,738	4,400	750
PDAC 8	1,950,197	574	316	9
PDAC 9	3,853,773	19,572	12,563	1,731
PDAC 10	2,748,850	28,225	18,767	2,489
PDAC 11	2,848,599	23,163	14,634	2,589
PDAC 12	3,723,326	21,243	12,940	2,122
PDAC 13	1,834,743	24,549	15,342	3,150
PDAC 14	2,481,332	25,650	18,016	2,564
PDAC 15	1,752,081	38,514	25,899	5,210
Normal Pancreas 1	1,196,372	908	284	0
Normal Pancreas 2	975,676	1,043	303	0
Lung Cancer 1	1,549,237	28,658	18,751	4,417
Lung Cancer 2	13,829,845	33,030	26,143	2,555
Kidney Cancer 1	2,104,859	10,814	6,505	1,501
Kidney Cancer 2	4,753,409	5,025	2,739	625
Ovarian Cancer 1	12,596,542	26,658	14,513	3,074
Ovarian Cancer 2	7,290,000	4,089	2,058	403
Prostate Cancer 1	3,376,849	43,730	22,244	9,793
Prostate Cancer 2	12,052,244	23,947	14,201	3,209
Prostate Cancer 3	3,631,148	21,411	12,390	2,804
Normal Fetal Brain	384,453	2,843	1,516	3
Normal Brain	371,161	5,184	2,573	3
Normal Colon	183,855	13,059	7,229	5
Normal Fetal Liver	147,977	11,218	5,879	7
Normal Liver	117,976	7,968	3,730	25
Normal Lung	208,089	15,027	7,857	5
Normal Kidney	144,173	15,218	8,094	7
Normal Placenta	207,929	13,990	7,815	0
Normal Prostate	263,406	8,409	2,228	19
Normal Uterus	477,480	2,702	1,395	2

Table S4: Mouse Satellite Correlated Genes

Annotated transcripts with high linearity with major satellite repeat (R > 0.85) and classification using DAVID NIH Database.

GENE NAME	Neural	Germ/Stem Cell	Zinc Finger Domain
1190005F20Rik	x		X
1200014J11Rik		x	
2510002D24Rik			
2810021B07Rik	x	x	
2810408P10Rik	x		X
4832428D23Rik			
4930486G11Rik	x		
4930555G01Rik	x	x	
4932411N23Rik		x	
6330439K17Rik	x		
9330155M09Rik	x		
9630025I21Rik			x
9930013L23Rik	x		
A130023I24Rik			
AA388235	x	x	
Acer1			
Acss3			
Adamts6			
Adra1a	x		
Aff2	x		
Agbl3	x		
Ap4e1	x		
Aplf	x		
Apol8			
Aqp7		x	
Armc9	x	x	
Cdnf			
Astn2	x		
Atg16l1	x	x	
B020018G12Rik			
Bbs1		x	
Btbd7	x	x	
Cadm3	x		
Cage1	x	x	

Camk2a	x		
Ccdc114		x	
Ccdc132	x	x	
Cckar	x		
Cenpc1	x		
Chst3	x		
Clasp2	x		
Clec4a2	x		
Cplx2	x		
Crnkl1		x	
Crx	x	x	
Csmd3	x	x	
Cx3cr1	x		
Dclre1b	x	x	
Defb42	x		
Dnajb14		x	
Dnm3	x		
Dok6			
Dub1	x		
Eif2ak2			
Eif4e3		x	
Ermp1		x	
Fam110b	x		
Fgd4	x		
Gabra2	x		
Galnt13	x		
Gcm1		x	
Gga2	x		
Gje1	x		
Gk5		x	
Glipr1l2		x	
Glrx2	x	x	
Gm13242			x
Gm4841			
Gm5094			
Gm5546	x		
Gm7120	x		
Shisa6			
Gpr98	x		
Gria3	x		
Grk1	x		

Grk4	x	x	
Hecw1	x		
Hspbap1			
Ikbkb	x		
Il17rd		x	x
Il18rap			
Il21	x		
Il2ra			
Cxcr2		x	
Irf9			
Itga4	x	x	
Kcnj10	x	x	
Kcnmb1	x		
Khdrbs2	x	x	
Kl	x		
L1cam	x		
L1td1 (Line-1)		x	
Lag3	x		
Lefty2		x	
Lipi	x		
Lmf1			
Lyrm1		x	
Plin3			
Malt1	x		
Me3	x	x	
Med16	x		
Mrgprb1			
Mrpl13		x	
Ms4a7			
Msx2			
Mtif2		x	
Mto1			
Myh8	x		
Mzf1	x		x
Nhedc1		x	
Nkain3	x		
Nol3	x		
Npy1r	x		
Nsun7		x	
Obox3		x	
Olfr628			

Onecut3	x		
Otx1	x		
P4ha2			
Pard3b	x		
Pcdh10	x		
Pde4a	x		
Cdk14	x	x	
Phka1	x		
Piga	x		
Pik3r4	x		
Pkn1	x		
Plcxd3	x		
Polq	x	x	
Prdm8	x		x
Prkcb	x		
Prkg1	x		
Prmt10			
Prosapip1			
Psg28	x		
Psmb11			
Rab8b	x		
Rag2			
Rbm34	x		
Rfesd	x		
Rgs9	x		
Rgs9bp	x		
Rhbdl2			
Rnf150	x		
Scd4			
Serpina3a		x	
Sfrp4	x		
Six2			
Skint4			
Slc17a8	x		
Slc22a9			
Slc25a40	x		
Slc26a8	x		
Slc2a4	x		
Slc4a1ap	x		
Slc6a2	x		
Slc6a5	x		

Slco1c1	x		
Slfn8			
Slfn9			
Snai2			x
Spaca1	x	x	
Spc25	x		
Sfrs13b			
St8sia1	x		
Sult3a1	x		
Supt7l	x	x	
Syt15			
Tacr1	x		
Tbx5	x		
Tctn2	x		
Prss43			
Thada	x	x	
Tigd3	x	x	
Tigd4			
Tmem91	x		
Topbp1	x		
Tox4		x	
Traf5			
Trhr			
Trpm3	x		
Tspan18			
Ttc21b	x		
Uevld	x		
Ulk3	x		
Vmn1r58			
Vmn2r55			
Xpo7	x	x	
Zfp353			x
Zfp459			x
Zfp473		x	x
Zfp488	x		x
Zfp560	x		x
Zfp874	x		x
Zfyve28			
Zmat1		X	x
Zmat3	x		x
Zmym5	x		

uc007ais.1			
uc007bjl.1			
uc007bvm.1			
uc007das.1			
uc007epk.1			
uc007epl.1			
uc007etc.1			
uc007fow.1			
uc007fye.1			
uc007gji.1			
uc007gyq.1			
uc007ilo.1			
uc007kpk.1			
uc007kpl.1			
uc007ksq.1			
uc007ncm.1			
uc007oee.1			
uc007okg.1			
uc007pip.1			
uc007qaj.1			
uc007qgo.1			
uc007qwh.1			
uc007rhg.1			
uc007ric.1			
uc007rjz.1			
uc007tqu.1			
uc007vdl.1			
uc007vxz.1			
uc007wsb.1			
uc007ydk.1			
uc008aec.1			
uc008aki.1			
uc008dlt.1			
uc008dwg.1			
uc008dwr.1			
uc008dyu.1			
uc008ekc.1			
uc008elg.1			
uc008ewf.1			
uc008fxu.1			
uc008ghy.1			

uc008hdz.1			
uc008hes.1			
uc008hln.1			
uc008hym.1			
uc008hyp.1			
uc008ibk.1			
uc008ivz.1			
uc008jjy.1			
uc008kvl.1			
uc008laa.1			
uc008lfo.1			
uc008lqq.1			
uc008lsd.1			
uc008mru.1			
uc008msi.1			
uc008msj.1			
uc008ofl.1			
uc008ogy.1			
uc008ohe.1			
uc008ovz.1			
uc008oxq.1			
uc008oxr.1			
uc008pbs.1			
uc008pdh.1			
uc008pwt.1			
uc008qgg.1			
uc008qnj.1			
uc008rib.1			
uc008rks.1			
uc008tmh.1			
uc008vsq.1			
uc008vss.1			
uc008vsw.1			
uc008wgl.1			
uc008wpr.1			
uc008yil.1			
uc008ype.1			
uc008ypk.1			
uc008zfa.1			
uc009auf.1			
uc009biv.1			

uc009bte.1			
uc009eds.1			
uc009fxk.1			
uc009gaq.1			
uc009ify.1			
uc009lfw.1			
uc009lop.1			
uc009lxi.1			
uc009mey.1			
uc009mgu.1			
uc009pbg.1			
uc009qbe.1			
uc009tki.1			
uc009tzq.1			
uc009tzr.1			
uc009vam.1			
uc009vfd.1			

Table S5: Human Satellite Correlated Genes
Annotated human transcripts with high linearity with ALR satellite (R > 0.85) and classification using DAVID NIH Database.

GENE NAME	Neural	Germ/Stem Cell	Zinc Finger Domain
A2ML1			
ABCA9	x	x	
ACADSB			
ACBD7	x		
ADAMTSL3	x		
ALG11			
ANGEL2	x		
ANKRD20A1		x	
AP1S3	x		
APOL4			
APOL6			
ATP10B	x		
BNC1			x
C11ORF72			
C11ORF74			
C12ORF5			
C13ORF29	x		
C15ORF2			
C15ORF28		x	
C17ORF77			
C1ORF130	x		
C1ORF69			
C1ORF84	x	x	
C21ORF82	x		
C3ORF20	x	x	
C6ORF170	x	x	
C7ORF44		x	
C7ORF46	x		
C8ORF12			
C9ORF68	x		
CAGE1		x	
CCBP2			
CCDC122			
CCDC52	x		
CD3EAP	x	x	
CDON	x		
CENPM	x		
CES3		x	
CES7	x		
CHRM5	x	x	
CLCC1	x		
COX18	x		

CPM			
CPSF2	x	x	
CYP46A1	x		
DBF4B			x
DCHS2	x		
DDO	x		
DHRS4L2			
DKFZP434L187		x	
DKFZP779L1853			
DNAH5	x	x	
DNAH8		x	
DSG3			
DUSP19		x	
DZIP3	x	x	x
EEF2K			
F2RL3			
FAM111B		x	
FAM122C		x	
FAM22G			
FAM75A2			
FAM83D			
FAT3	x		
FBXO15	x	x	
FBXW10		x	
FCF1	x		
FER			
FGF5	x		
FLJ11292			
FLJ41649			
FLJ43763		x	
FUT1	x		
GALNT13	x	x	
GBP4	x	x	
GK5	x		
GLIPR1L2	x	x	
GPR110			
GPR157			
GTPBP10	x		
GTSE1	x		
GUSBP1			
HERC4	x		
HESRG		x	
HIF3A	x	x	
HMGA2			
HRH4			
HUNK			
HYDIN	x	x	

IL12RB1			
IPO9	x		
KCTD18	x	x	
KIAA1245			
KIAA1257	x	x	
KIAA1328	x		
KIR3DX1			
LEPRE1	x	x	
LOC147804			
LOC349196			
LOC440313			
LOC441242		x	
LOC441426			
LOC642980			
LOC643406	x	x	
LOC649305	x		
LOC91948			
LRRC2	x		
LTV1			
LYRM2		x	
LYRM7	x		
MCFD2	x		
MED18			
MORC4		x	x
MSH5		x	
MTBP		x	
MX2	x		
MYH1			
MYO3B			
MYOM3	x	x	
NBPF1			
NEB	x		
NHEDC1	x	x	
NIPSNAP3B			
NME7	x		
NMNAT1			
NUP43		x	
ODF2L	x		
OR11H1	x		
OR11H12	x		
OR4F16	x		
OR4K15	x		
OR7D2	x		
OR7E156P	x		
ORC6L		x	
PCBD2			
PDDC1	x		

PGPEP1	x		
PHACTR4	x	x	
PHTF1			
PLA2G2D			
PLEKHA5	x		
PRKRIR	x		x
PRND	x	x	
PXMP4	x		
QTRTD1		x	
RASGRP3	x		x
REXO1L1			
RGR	x		
RNF125		x	x
SIGLEC10			
SIGLEC8			
SIRPB1	x		
SLC13A2			
SLC14A2		x	
SLC16A12			
SLC19A3	x		
SLC1A6	x	x	
SLC27A1			
SLC31A1	x		
SMU1		x	
SP100			
STRC			
STX17	x	x	
TAOK1	x		
TCL6			
TEX9	x		
TGFB2			
TIGD1	x		
TNFRSF19	x		
TRIM43			x
TRPM3	x		
TTN	x	x	
ULBP1	x	x	
USPL1	x	x	
UTP14C			
WDR17	x		
WDR31		x	
XKR9			
XRCC2	x	x	
ZFYVE20	x		x
ZMYM1		x	x
ZMYND17	x		x
ZNF100		x	x

ZNF192		x	x
ZNF208			x
ZNF273			x
ZNF320	x		x
ZNF331		x	x
ZNF37A			x
ZNF383		x	x
ZNF431	x	x	x
ZNF445	x		x
ZNF471	x		x
ZNF480	x	x	x
ZNF490	x	x	x
ZNF492	x		x
ZNF493	x		x
ZNF528	x	x	x
ZNF562			x
ZNF621			x
ZNF623	x		x
ZNF667			x
ZNF670			x
ZNF7			x
ZNF720			x
ZNF804B	x		x
BC029464			
BC082237			
BC050580			
BC039319			
AK096834			
BC042893			
BC043508			
HBET1			
NR_003246			
LOC643079			
BC040190			
AK095450			
BC036442			
DKFZP761G18121			
AK092337			
KIAA0379			
FLJ44076			
AX748237			
AX747345			
AX747165			
CR627148			
UNQ2963			
DKFZP667M2411			
AK125319			

AK125996			
AK026805			
AK129982			
CR592614			
AK095077			
BC035989			
CR623134			
AK026100			
RP1-140A9.6			
AX747405			
NR_002828			
NR_003130			
BIRC4BP			
AK054836			
AX747417			
AY314745			
NR_001318			
AX747586			
AK125128			
AK055694			
BC035084			
WUGSC:H_DJ0855D21.2			
CR596262			
AX746734			
AK024378			
BC037952			
BC041998			
BC008050			
NR_003133			
AX748369			
BC043541			
AK131347			
FLJ00140			
CR620525			
AX748243			
AX747639			
AX746484			
CR605783			
AK097143			
BC052952			
AK124179			
FLJ16008			
BC073807			
BC015784			
CR592225			
BC031280			
DKFZP686F19123			

AX747440			
AK096469			
AK124893			
AX747721			
AK123584			
NR_003263			
DKFZP762C213			
BC094791			
CR627394			
AK124673			
NR_002910			
FRABIN			
BC069727			
BC037884			
BX648696			
CR627383			
BC034569			
AX747308			
AK123585			
BC011779			
DKFZP686H1615			
BC070093			
BX537874			
AX748226			
CR598144			
BC040189			
AL832479			
NR_002939			
AL833449			
BC047600			
KIAA1031			
AK095766			
AL832786			
BC035181			
NR_002220			
DQ596646			
NM_001001704			
AL832797			
AK129672			
AK123838			
AX746771			
C20ORF38			
AX746989			
LOC285382			
MGC102966			
AK124194			
FLJ45337			

AK126334			
AK057596			
NR_003128			
AK096077			
DERP7			
AK098126			
BC033330			
BC029555			
LOC129881			
AK097527			
BX648961			
AK096499			
AK097777			
AK091028			
FLJ37953			
PTPN1L			
AK096196			
AK056351			
AX746750			
LOC440053			
BC068605			
UNQ9369			
PFDN6L			
AK125042			
AK125489			
BC013681			
AK056866			
AX747590			
AX746620			
FLJ00310			
NM_001042703			
AK094618			
AX748002			
BC041646			
AJ617629			
AL833139			
AK097428			
AK056105			
MGC13098			
AK127557			
KIAA1456			
BC069809			
LOC441108			
NM_001039909			
AK096291			
BX537710			
BC041449			

NR_002836			
CR598129			
BC035112			
CR613732			
DQ597733			
AX747172			
AK128266			
TCAM-1			
BC050344			
BC047380			
AL832439			
BC042121			
BC041426			
C15ORF20			
AK125310			
DKFZP434P055			
KIAA0010			
COX18HS			
BC038578			
AY314748			
AK023134			
AK131313			
BC041865			
AX746851			
LOC606495			
AK127238			
LOC441282			
BOZF1			
AK026825			
AK128305			
AL713649			
DQ573949			
AK091996			
CR606964			
HSKRP1			
AX747556			
NR_003266			
CR749689			
BC049371			
AX747988			
FLJ35848			
WHDC1L1			
AK126491			
AK024841			
AX746688			
FLJ37357			
FLJ44955			

BC040631			
CR627135			
DKFZP451M2119			
CR627206			
AK127460			
BC019672			
HERV-H-HHLA1_FUSION			
AK057632			
FLJ00264			
NY-REN-7			
AK125288			
AF086203			
LOC94431			
BC043415			
AK098333			
BC042588			
AX747864			
AY314747			
AK128216			
BC044257			
AX747062			
BX649144			
AL137270			
PP8961			
AK056558			
AK094845			
AX747742			
AK095981			
CTRP6			
NR_002821			
AX746880			
AK125817			
AK056417			
AK026469			
AK090984			
AK131520			
AL833246			
AK125832			
BC041455			
AF380582			
AX747658			
AX721193			
BC047626			
FLJ44060			
KIAA0982			
AK093513			
BC038431			

BX161428			
DKFZP686O248			
AK096335			
BX640887			
BC009626			
AY338954			
BC036412			
NM_001001681			
AK056892			
DQ573361			
BC041466			
NR_002210			
FLJ33706			
KIAA1767			
MBL1P1			
BC071776			
AK127888			
NR_002943			
AX747340			
LOC401252			
AX746585			
AK091594			
AK096412			
FLJ34047			
AX747756			
BC090058			
CR611653			
AL137733			
BX537706			
NR_001565			
MGC4836			
MGC29891			
AK098240			
AX748249			
C1ORF140			
AK055868			
BC122562			
BC041363			
BC047625			
BC021741			
AK056524			
BX647358			
AK023515			
AK125311			
AK123891			
LOC339809			
AK128523			

AK094859			
PJCG6			
AX748371			
UNQ3037			
AK054880			
AK094224			
AL833510			
KENAE1			
BC012110			
BC052779			
AK097893			
BC105727			
AK091527			
WBSCR23			
BC043378			
AK056246			
LOC401898			
AK023856			
UNQ1849			
BC048997			
FLJ36492			
KIAA2023			
AK054869			
CR749689			
BC029555			
AK024378			
NR_002821			
DKFZP686F19123			

Table S6: Categorization of annotated satellite correlated genes
Satellite correlated genes were classified the using DAVID gene ontology program. Shown are the total numbers of genes in the neural, germ/stem cell, and zinc finger categories. Percentage of total genes mapped by the DAVID gene ontology program shown in parentheses.

	Neural	Germ/Stem Cell	Zinc Finger Domain
Mouse SCGs	120 (63%)	50 (26%)	16 (8%)
Human SCGs	101 (49%)	63 (31%)	35 (17%)

Table S7: Analysis of neuroendocrine markers in human PDAC with high and low ALR satellite levels. Human PDACs were split into low and high total satellite expression using the median satellite expression. Average expression of each gene in low and high satellite human pancreatic tumors shown. Fold change and t-test q-value between high and low satellite PDAC samples is shown. Positive fold change indicates expression higher in high satellite PDAC while negative fold change indicates expression higher in low satellite PDAC.

		AVERAGE EXPRESSION (tpm)			
Human Gene Symbol	**Human Gene Name**	**PDAC LOW ALR**	**PDAC HIGH ALR**	**FOLD**	**q-value (FDR)**
SVOP	SV2 related	15.60	58.11	3.73	0.015
SYN2	Synapsin 2	5.96	18.17	3.05	0.015
CHGB	Chromagranin B	2.95	10.08	3.41	0.272
SYN1	Synapsin 1	1.16	3.20	2.76	0.272
SCG2	Chromagranin C	3.10	8.93	2.88	0.272
SLC6A4	Serotonin	1.27	3.81	3.00	0.284
VAMP1	Synaptobrevin	15.73	21.32	1.35	0.424
NCAM1	CD56	52.74	21.25	-2.48	0.424
ASCL1	ASCL1	20.58	14.12	-1.46	0.536
ENO2	NSE	40.82	33.51	-1.22	0.536
CHGA	Chromagranin A	11.97	9.19	-1.30	0.563
SYP	Synaptophysin	27.92	25.61	-1.09	0.563

Table S8: List of candidate histone modifiers evaluated in mouse and human samples with high and low satellite levels. A list of histone modifiers was complied from three recent reviews of epigenetic modifiers (*S9-S11*).

Count	Human Gene Name	Mouse Gene Name	Other Name	FUNC	Reference
1	CLOCK	Clock	KAT13D	Acetyltransferase	1
2	CREBBP	Crebbp	KAT3A	Acetyltransferase	1
3	ELP3	Elp3	KAT9	Acetyltransferase	1
4	EP300	Ep300	KAT3B	Acetyltransferase	1
5	Gtf3c4	Gtf3c4	KAT12	Acetyltransferase	1
6	HAT1	Hat1	KAT1	Acetyltransferase	1
7	GCN5L2	Kat2a	KAT2A	Acetyltransferase	1
8	PCAF	Kat2b	KAT2B	Acetyltransferase	1
9	HTATIP	Kat5	KAT5	Acetyltransferase	1
10	MYBBP1A	Mybbp1a	KAT13C	Acetyltransferase	1
11	Myst1	Myst1	KAT8	Acetyltransferase	1
12	Myst2	Myst2	KAT7	Acetyltransferase	1
13	Myst3	Myst3	KAT6A	Acetyltransferase	1
14	Myst4	Myst4	KAT6B	Acetyltransferase	1
15	NCOA1	Ncoa1	KAT13A	Acetyltransferase	1
16	NCOA3	Ncoa3	KAT13B	Acetyltransferase	1
17	TAF1	Taf1	KAT4	Acetyltransferase	1
18	EGLN1	Egln1		Demethylase	2
19	EGLN2	Egln2		Demethylase	2
20	EGLN3	Egln3		Demethylase	2
21	HIF1AN	Hif1an		Demethylase	2
22	HR	Hr		Demethylase	2
23	HSPBAP1	Hspbap1		Demethylase	2
24	JARID2	Jarid2		Demethylase	2
25	JHDM1D	Jhdm1d		Demethylase	2
26	JMJD1C	Jmjd1c		Demethylase	2
27	JMJD4	Jmjd4		Demethylase	2
28	JMJD6	Jmjd6		Demethylase	2
29	LOC339123	Jmjd8		Demethylase	2
30	LSD	Kdm1	LSD1	Demethylase	2
31	FBXL11	Kdm2a		Demethylase	2
32	FBXL10	Kdm2b		Demethylase	2
33	JMJD1A	Kdm3a		Demethylase	2
34	JMJD1B	Kdm3b		Demethylase	2
35	JMJD2A	Kdm4a		Demethylase	2
36	JMJD2B	Kdm4b		Demethylase	2
37	JMJD2C	Kdm4c		Demethylase	2
38	JMJD2D	Kdm4d		Demethylase	2
39	JARID1B	Kdm5b	PLU1	Demethylase	2
40	JARID1C	Kdm5c	SMCX	Demethylase	2
41	JARID1D	Kdm5d	SMCY	Demethylase	2
42	UTX	Kdm6a		Demethylase	2

43	JMJD3	Kdm6b		Demethylase	2
44	PHF2	Phf2		Demethylase	2
45	PHF8	Phf8		Demethylase	2
46	PLA2G4B	Pla2g4b		Demethylase	2
47	HDAC1	Hdac1		HDAC	2
48	HDAC10	Hdac10		HDAC	2
49	HDAC11	Hdac11		HDAC	2
50	HDAC2	Hdac2		HDAC	2
51	HDAC3	Hdac3		HDAC	2
52	HDAC4	Hdac4		HDAC	2
53	HDAC5	Hdac5		HDAC	2
54	HDAC6	Hdac6		HDAC	2
55	HDAC7A	Hdac7		HDAC	2
56	HDAC8	Hdac8		HDAC	2
57	HDAC9	Hdac9		HDAC	2
58	ASH1L	Ash1l		Methyltransferase	3
59	EHMT1	Ehmt1	GLP1	Methyltransferase	3
60	EHMT2	Ehmt2	G9a	Methyltransferase	3
61	EZH1	Ezh1		Methyltransferase	3
62	EZH2	Ezh2		Methyltransferase	3
63	MDS1	Mds1	PRDM3	Methyltransferase	3
64	MLL	Mll1	ALL-1	Methyltransferase	3
65	MLL2	Mll2		Methyltransferase	3
66	MLL3	Mll3		Methyltransferase	3
67	MLL5	Mll5	KMT2E	Methyltransferase	3
68	NSD1	Nsd1		Methyltransferase	3
69	PRDM1	Prdm1		Methyltransferase	3
70	PRDM13	Prdm13		Methyltransferase	3
71	PRDM14	Prdm14		Methyltransferase	3
72	PRDM15	Prdm15		Methyltransferase	3
73	PRDM16	Prdm16		Methyltransferase	3
74	PRDM2	Prdm2		Methyltransferase	3
75	PRDM5	Prdm5		Methyltransferase	3
76	PRDM6	Prdm6		Methyltransferase	3
77	PRDM8	Prdm8		Methyltransferase	3
78	PRDM9	Prdm9		Methyltransferase	3
79	KIAA0339	Setd1a	SET1A	Methyltransferase	3
80	KIAA1076	Setd1b	SET1B	Methyltransferase	3
81	SETD2	Setd2	HYPB	Methyltransferase	3
82	SETD3	Setd3		Methyltransferase	3
83	SETD6	Setd6		Methyltransferase	3
84	SETD8	Setd8	PR-SET7	Methyltransferase	3
85	SETDB1	Setdb1	ESET	Methyltransferase	3
86	SETDB2	Setdb2	CLL8	Methyltransferase	3
87	SETMAR	Setmar		Methyltransferase	3
88	SMYD1	Smyd1		Methyltransferase	3
89	SMYD2	Smyd2		Methyltransferase	3
90	SMYD3	Smyd3		Methyltransferase	3
91	SMYD4	Smyd4		Methyltransferase	3

92	SMYD5	Smyd5		Methyltransferase	3
93	SUV39H1	Suv39h1		Methyltransferase	3
94	SUV39H2	Suv39h2		Methyltransferase	3
95	SUV420H1	Suv420h1		Methyltransferase	3
96	SUV420H2	Suv420h2		Methyltransferase	3
97	MLL4	Wbp7	MLL4	Methyltransferase	3
98	WHSC1	Whsc1	NSD2	Methyltransferase	3
99	WHSC1L1	Whsc1l1	NSD3	Methyltransferase	3

Table S9: List of candidate histone modifiers evaluated in mouse PDAC tumors with high and low satellite levels. Mouse PDACs were split into low and high satellite expression using the median satellite expression. Average expression of each gene in low and high satellite mouse pancreatic tumors shown. Fold change and t-test q-value between high and low satellite PDAC samples is shown. Positive fold change indicates expression higher in high satellite PDAC while negative fold change indicates expression higher in low satellite PDAC.

		AVERAGE EXPRESSION (tpm)			
Mouse Gene Name	**Function**	**PDAC LOW SAT**	**PDAC HIGH SAT**	**FOLD**	**q-value (FDR)**
Hspbap1	Demethylase	20.5	104.8	5.11	0.019
Kdm3b	Demethylase	73.7	34.2	-2.16	0.020
Myst4	Acetyltransferase	35.7	92.2	2.58	0.021
Jmjd4	Demethylase	3.9	12.1	3.10	0.021
Hdac9	HDAC	26.7	115.9	4.34	0.021
Kdm5d	Demethylase	9.7	95.2	9.84	0.021
Kdm6b	Demethylase	51.3	83.4	1.62	0.057
Hdac3	HDAC	121.7	57.2	-2.13	0.057
Hdac4	HDAC	33.1	16.6	-2.00	0.057

Table S10: List of candidate histone modifiers evaluated in human PDAC tumors with high and low satellite levels. Human PDACs were split into low and high total satellite expression using the median ALR satellite expression. Average expression of each gene in low and high satellite human pancreatic tumors shown. Fold change and t-test q-value between high and low satellite PDAC samples is shown. Positive fold change indicates expression higher in high satellite PDAC while negative fold change indicates expression higher in low satellite PDAC.

		AVERAGE EXPRESSION (tpm)			
Human Gene Name	**Function**	**PDAC LOW ALR**	**PDAC HIGH ALR**	**FOLD**	**q-value (FDR)**
SETMAR	Methyltransferase	12.6	35.3	2.8	0.057
SMYD4	Methyltransferase	143.7	339.7	2.4	0.057
PRDM6	Methyltransferase	1.9	5.9	3.1	0.057
SMYD1	Methyltransferase	3.3	12.6	3.8	0.057

SUPPLEMENTARY ONLINE MATERIALS REFERENCES

S1. N. Bardeesy *et al.*, *Proc Natl Acad Sci U S A* **103**, 5947 (2006).
S2. A. J. Aguirre *et al.*, *Genes Dev* **17**, 3112 (2003).
S3. D. Lipson *et al.*, *Nat Biotechnol* **27**, 652 (2009).
S4. D. Pushkarev, N. F. Neff, S. R. Quake, *Nat Biotech* **27**, 847 (2009).
S5. J. Jurka *et al.*, *Cytogenet Genome Res* **110**, 462 (2005).
S6. G. Dennis, Jr. *et al.*, *Genome Biol* **4**, P3 (2003).
S7. W. Huang da, B. T. Sherman, R. A. Lempicki, *Nat Protoc* **4**, 44 (2009).
S8. J. D. Storey, R. Tibshirani, *Methods Mol Biol* **224**, 149 (2003).
S9. M. Albert, K. Helin, *Semin Cell Dev Biol* **21**, 209 (2010).
S10. C. D. Allis *et al.*, *Cell* **131**, 633 (2007).
S11. P. A. Cloos, J. Christensen, K. Agger, K. Helin, *Genes Dev* **22**, 1115 (2008).
S12. F. Li, L. Sonbuchner, S. A. Kyes, C. Epp, K. W. Deitsch, *J Biol Chem* **283**, 5692 (2008).
S13. J. H. Martens *et al.*, *EMBO J* **24**, 800 (2005).