Supplemental Figure Legends

Supplemental Figure S1, related to Fig. 1:

A: SMAD3 T179 phosphorylation is not required for TGF-β-induced transcriptional activity. HeLa cells stably expressing various SMAD3 constructs were transfected with SMAD3-dependent $(CAGA)_{12}$-Luc reporter, and luciferase activities were measured after treated with ± TGF-β for overnight. Data are represented as mean ± S.D. (n=3).

B: EGF enhanced TGF-β-induced Vimentin, TWIST and MMP2 mRNA expression in HeLa cells, but inhibited TGF-β-mediated transcriptional response on p21 and c-MYC expression. qRT-PCR analyses of gene expression in HeLa cells stably expressing various SMAD3 constructs. Relative levels of expression were normalized to untreated cells. Data are represented as mean ± S.D. (n=3), statistically significant difference between SMAD3 vector expressing cells and control shNS cells were indicated, * $p<0.02$. TGF-β treated samples were the same as in Figure 1C.

C: EGF accelerates TGF-β-induced EMT in MCF10A cells. Immunofluorescence staining of F-Actin (phalloidin) and E-Cadherin in MCF10A cells in the absence or presence of TGF-β or combined TGF-β and EGF treatment. Bar = 20 μm.

D: Expression of SMAD3 in MCF10A cells that stably expressing SMAD3 siRNA-resistant, Flag-tagged SMAD3 wild type (WT) or T179V (TV) construct. The endogenous SMAD3 was knockdown using siRNA against SMAD3.

E: Invasion of MCF10A cells across matrigel-coated membranes was assayed with untreated cells or cells treated with TGF-β and EGF (18 hr). Quantitation of invasion was performed by counting cells that invaded across matrigel membrane, data are represented as mean ± S.D. (n=3) statistically significant difference between siSmad3- and siNS-transfected cells were indicated, * $p<0.02$.

F: Xenograft assay for stable pools of HeLa cells in which tumor volumes were measured at 3 weeks after subcutaneous inoculation. Mean ± SD (12 tumors from 6 mice), and statistical significance (p value) between tumors from SMAD3WT and SMAD3TV cells are indicated.

Supplemental Figure S2, related to Fig. 2:

A: TGF-β and EGF-induced SMAD3 T179 phosphorylation can be inhibited by CDK inhibitor, flavopiridol and MEK1/2 inhibitor, U0126, respectively. AKT inhibitor VIII had little affect on SMAD3 T179 phosphorylation. HeLa cells were pretreated with chemical inhibitors for 4 hr prior to TGF-β or EGF stimulation for 2 hr and cells lysates were analyzed by Western blot using different phospho-SMAD3, and total SMAD3 antibodies.

B: TGF-β induced SMAD3 T179 and S423/S425 phosphorylation in HeLa cells.

C: TGF-α induced SMAD3 T179 phosphorylation in HeLa cells.

D: Schematic diagram of strategy used to identify nuclear proteins that interact with phosphorylated T179 of SMAD3.

E: Nuclear extracts of HeLa cells were subjected to affinity purification with agarose-beads coupled SMAD3 T179 or pT179 peptides. Proteins that pulled down on beads were run by a polyacrylamide gel and stained by Pageblue.

F: Co-localization of splicing factor SC35 with PCBP1 and SMAD3. GFP-SMAD3 or GFP-SMAD3(TV) transfected HeLa cells were treated with TGF-β and EGF (2 hr) and subjected to immunofluorescence using anti-SC35 and anti-PCBP1 antibodies. Bar = 10 μm.

G: Quantitation of proximal ligation assay signal of SMAD3 with PCBP1 in HeLa cells stably expressing various SMAD3 constructs. Mean ± SD (n=11) and statistically significant difference (p value) were shown in scatter dot plot.

H: AKT inhibitor VIII or EGFR inhibitor Afatinib inhibited PCBP1 nuclear localization. HeLa cells were pretreated with chemical inhibitors for 4 hr prior to combined TGF-β and EGF stimulation for 2 hr. Bar = 20 μm.

Supplemental Figure S3, related to Fig. 3:

A: Schematic diagram of structure of the CD44 gene (left) and alternatively spliced mRNA transcripts of mesenchymal isoform CD44 standard (CD44s) and epithelial isoform CD44E (right). The CD44s is encoded by standard exons: Exon 1-5, and Exon 16-19. Combinations of the variant exons, v1 to v10 (Exon 6-15), can be alternatively spliced between exon 5 and exon 16 to encode CD44 variable isoforms. CD44E contains variant exon v8, v9 and v10. Unlike the mouse gene, exon 6 (v1) of the human gene is not normally included in human CD44 mRNA. Primer pair used in qRT-PCR for pre-mRNA was noted in the left panel, primer pairs used for CD44 total RNA (black color), CD44s mRNA (green color) and CD44E mRNA (orange color) were noted in the right panel.

B: RT-PCR gel picture showing a switch in splicing of CD44 in MCF10A cells upon growth factor treatment for 72 hr. Primer pair used to amplify CD44s and CD44E at the same reaction was shown on the right.

C: qRT-PCR analysis of levels of CD44 isoforms in MCF10A cells. Relative levels of expression were normalized to untreated (Un) cells and shown as mean ± SD (n=3). Statistically significant difference ($p<0.02$) between treated and untreated samples were indicated by a black *, between combined TGF-β and EGF or TGF-α treated and TGF-β alone treated were indicated by a red *.

D: Chemical inhibitors of MEK1/2, CDK and AKT inhibited CD44 isoform switching in MCF10A cells. Relative levels of expression in cells treated with TGF-β and EGF were normalized to untreated (Un) cells, and shown as mean ± SD (n=3). Statistically significant difference ($p<0.02$) between treated and untreated samples were indicated by a black *, between inhibitor treated and control DMSO treated in the presence of TGF-β and EGF were indicated by a red *.

E: qRT-PCR analysis of CD44 variable exon ratios to standard exon 2 in HeLa cells stably expressing various vectors as indicated. Cells were treated with either TGF-β or EGF alone. Data were shown as in Fugure 3C.

F: Quantitation of relative fluorescence intensity of FACS analysis of cell surface expression of CD44 isoforms (related to Figure 3J). Data were shown as mean ± SD (n=3), statistically significant differences between treated and untreated samples were indicated.

Supplemental Figure S4, related to Fig. 4:

A. RNA samples from untreated (un) or DRB treated HeLa cells were analyzed by qRT-PCR using primer pair that detect transcription of CD44 pre-mRNA. The results showed that 4 hr of DRB treatment inhibited about 90% of transcription.

B: Schematic diagram showing the time course and procedure of the treatment (Top). DRB treatment synchronizes cells by inducing a transcription arrest, while removing of DRB then sets off a new wave of transcription. Examples of primers used to measure the levels of partially spliced mRNA and transcription of newly synthesized RNA were shown in the bottom panel. Primer set ExB and InB detects the transcription of pre-mRNA containing the exon B and intron B; primer set ExA/B and InB detects partially spliced RNA in which intron A has been spliced out but intron B remains unspliced. As the PoI II passes the Exon B, the ExB –InB transcript is made and accumulates in time, but as splicing catches up, Exon C would be spliced onto Exon B, while Intron B would be excised off. This would lead to an abrupt drop in the amount of the ExB –InB PCR amplification due to the removal of the InB primer binding sequence in Intron B, thus generating the peak of "relative RNA" levels in Figure 4. Likewise, the same peak of "relative RNA" levels would be generated with the ExA/B –InB pair after splice out intron A. However, if the splicing between alternate Exon B and C is blocked, the InB binding sequence will persist in the retained Intron B, causing the level of the and ExB –InB primary transcript to plateau, such as in the Figure 4B (middle panels, exon 12 and exon 14). On the other hand, if the splicing between alternate Exon A and B is blocked, the ExA/B primer binding sequence will fail to be reconstituted, causing the level of the ExA/B –InB Inx transcript to stay flat. After splice out intron B, both qRT-PCR products disappear.

C: qRT-PCR analyses of ESRP1 and ESRP2 expression induced by TGF-β treatment in HeLa cells. Relative levels of expression were normalized to untreated cells. Data are represented as mean ± S.D. (n=3), statistical significance between SMAD3 vector expressing cells and control shNS cells were indicated, * $p<0.02$.

Supplemental Figure S5, related to Fig. 5:

qRT-PCR analyses of the SMAD3 (left) or PCBP1 (right) RIP assays in pools of HeLa cells carrying shNS or shPCBP1 in the absence or presence of TGF-β and EGF treatment for 2 hr. Data are shown as mean ± SD (n=3). The black * denotes statistical significance ($p<0.02$) between treated and untreated samples.

Supplemental Figure S6, related to Fig. 6:

A: qRT-PCR analyses of CD44 isoforms in human mammary epithelial (HMLE) cells and oral cancer OSCC3 cells. Cells were treated with TGF-β and EGF together. Relative levels of expression were normalized to untreated (Un) cells and shown as mean ± SD (n=3). The black * denotes statistical significance ($p<0.02$) between treated and untreated samples.

B: qRT-PCR analyses of CD44 variable exon ratios to standard exon 2 in HMLE and OSCC3 cells. Cells were treated with TGF-β and EGF together. Data were shown as in Figure 3C. Noting difference of variable exon usage in HMLE and OSCC3 cells comparing to that of HeLa cells.

C: SMAD3 T179 phosphorylation in HMLE and OSCC3 cells upon T (TGF-β), E (EGF) or TE (TGF-β and EGF) treatment for 2 hr.

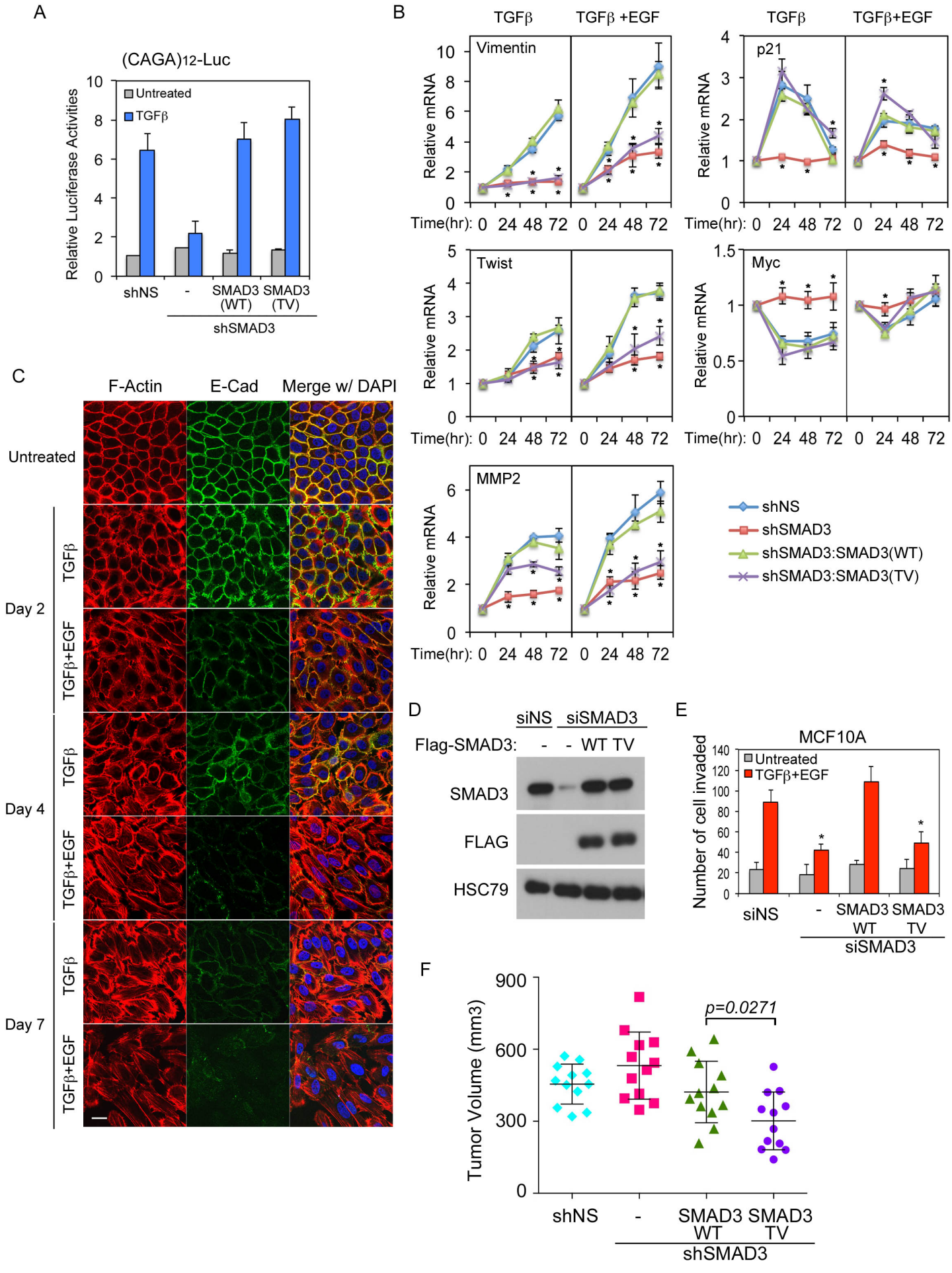
D: Western blot analyses of mesenchymal and metastasis associated protein, Vimentin, CTGF and MMP2, expression in HeLa cells expressing various SMAD3 vectors in the absence or presence of TGF-β and EGF treatment for 3 days. HSC70 was used as a loading control.

Supplemental Figure S7, related to Fig. 7:

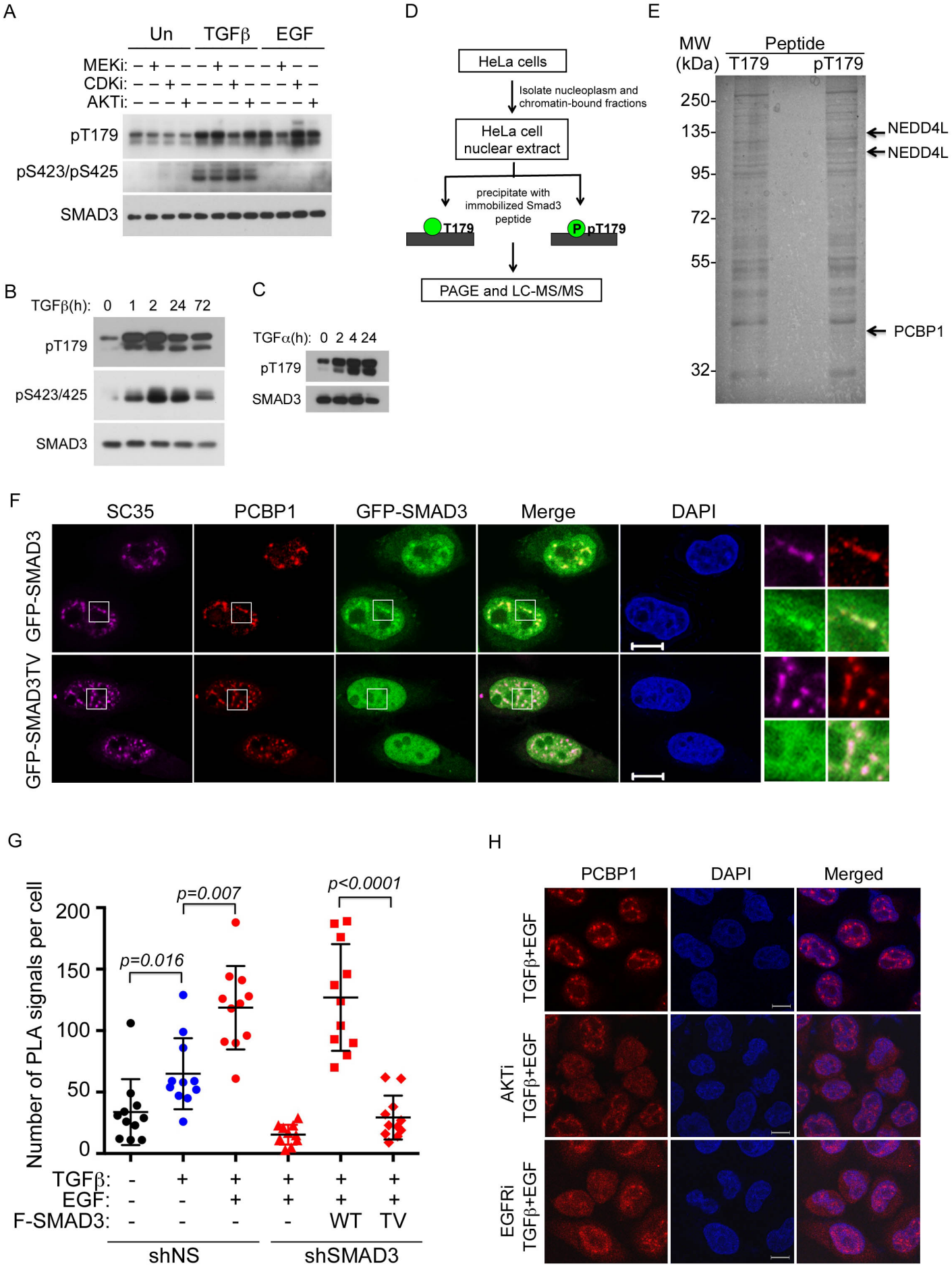
A: RT-PCR gel pictures showing that isoform switching of PDGFA, SLK, MAP3K7, FMNL3, and PTEN in HeLa cells treated with T (TGF-β), E (EGF) and TE (TGF-β and EGF) for 72 hr.

B: qRT-PCR validation of six SMAD3-independent SE events. Bars represent ratios of alternative exon/constant exon of individual genes. Cells were treated with indicated factors for 72 hr. Among them, alternative splicing of SREBF1 exon 2 was only dependent of EGF but not TGF-β. While TGF-β regulates alternative splicing of ARVCF exon 19, c22orf15 exon 9, ABl2 exon 2, IRF3 exon 2 and IRAK4 exon 4, the TGF-β-mediated effects did not require SMAD3.

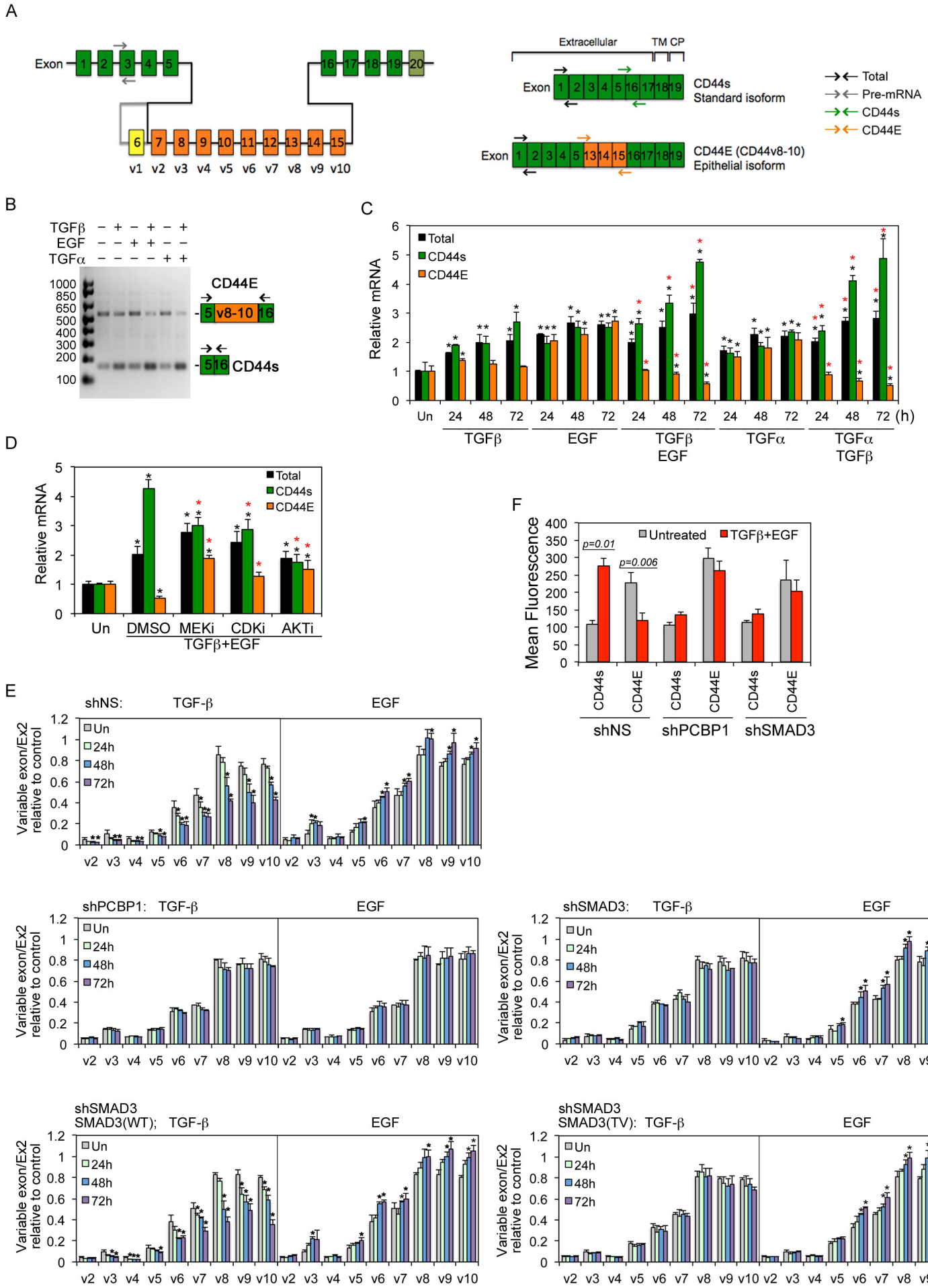
Supplemental Fig. S1, related to Fig. 1



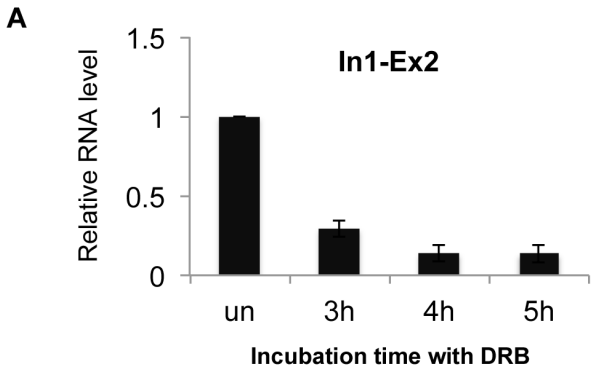
Supplemental Fig. S2, related to Fig. 2

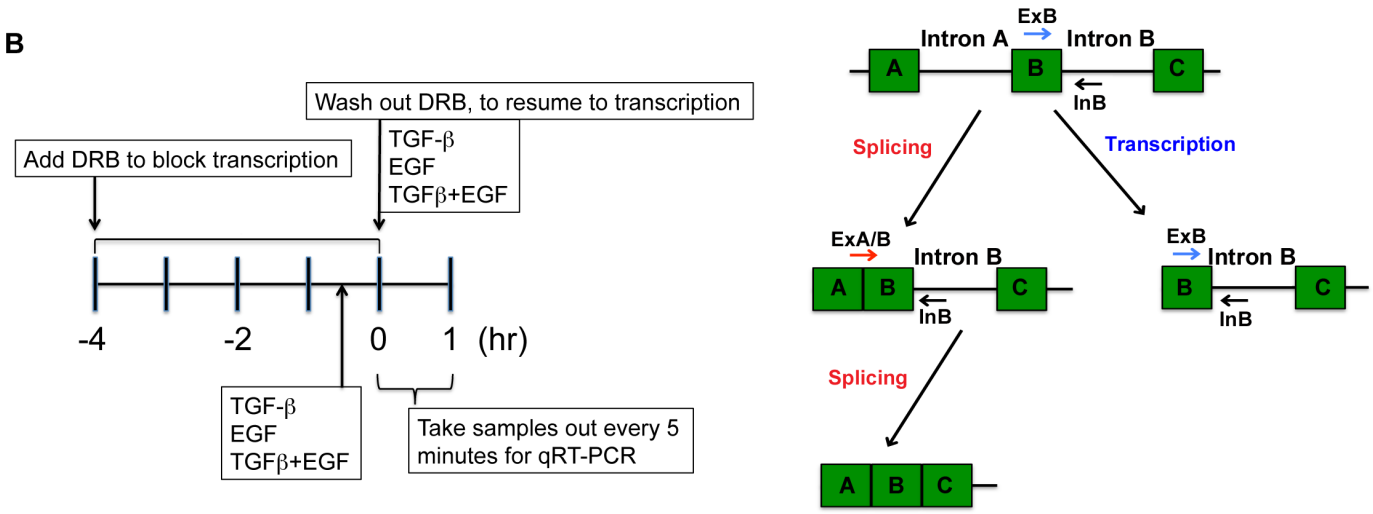


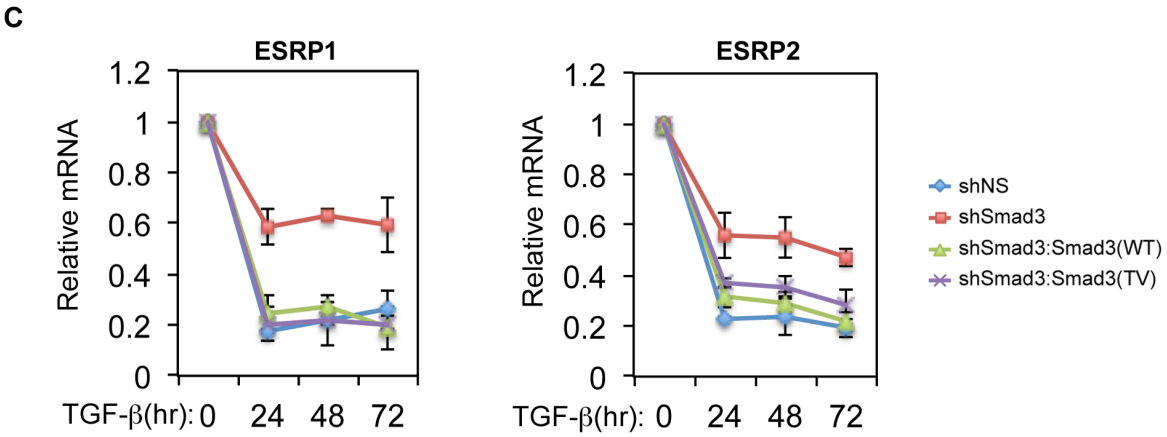
Supplemental Fig. S3, related to Fig. 3



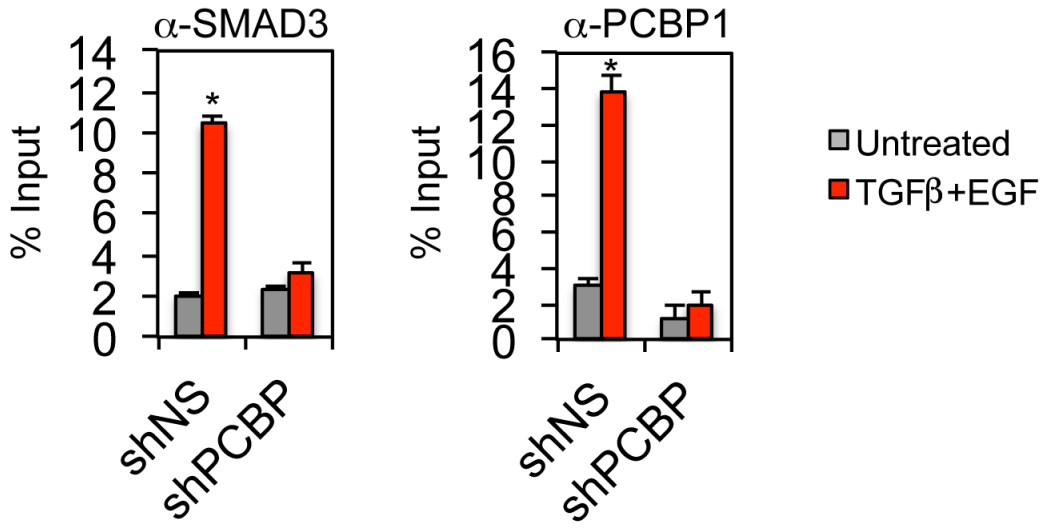
Supplemental Fig. S4, related to Fig. 4



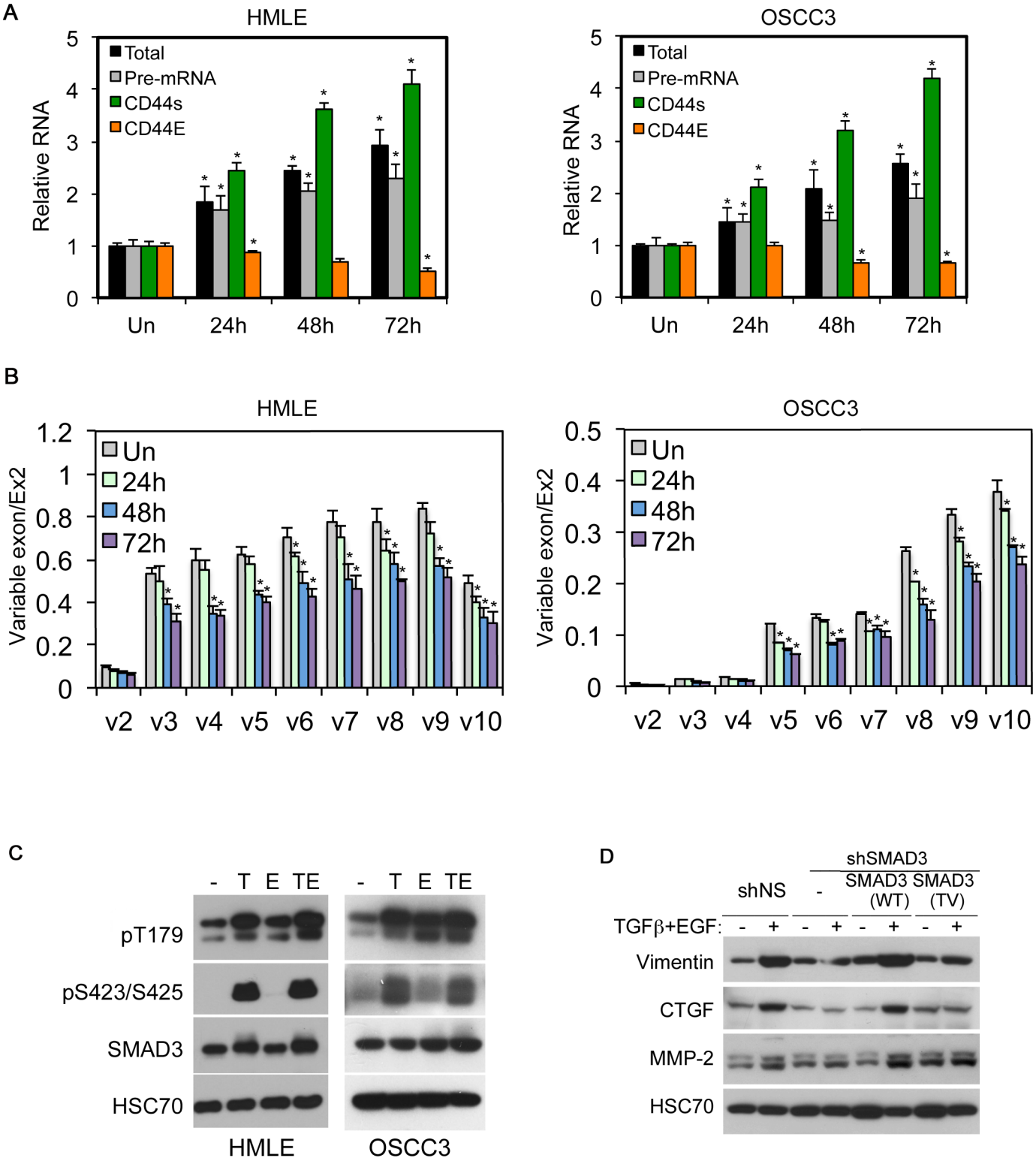




Supplemental Fig. S5, related to Fig 5

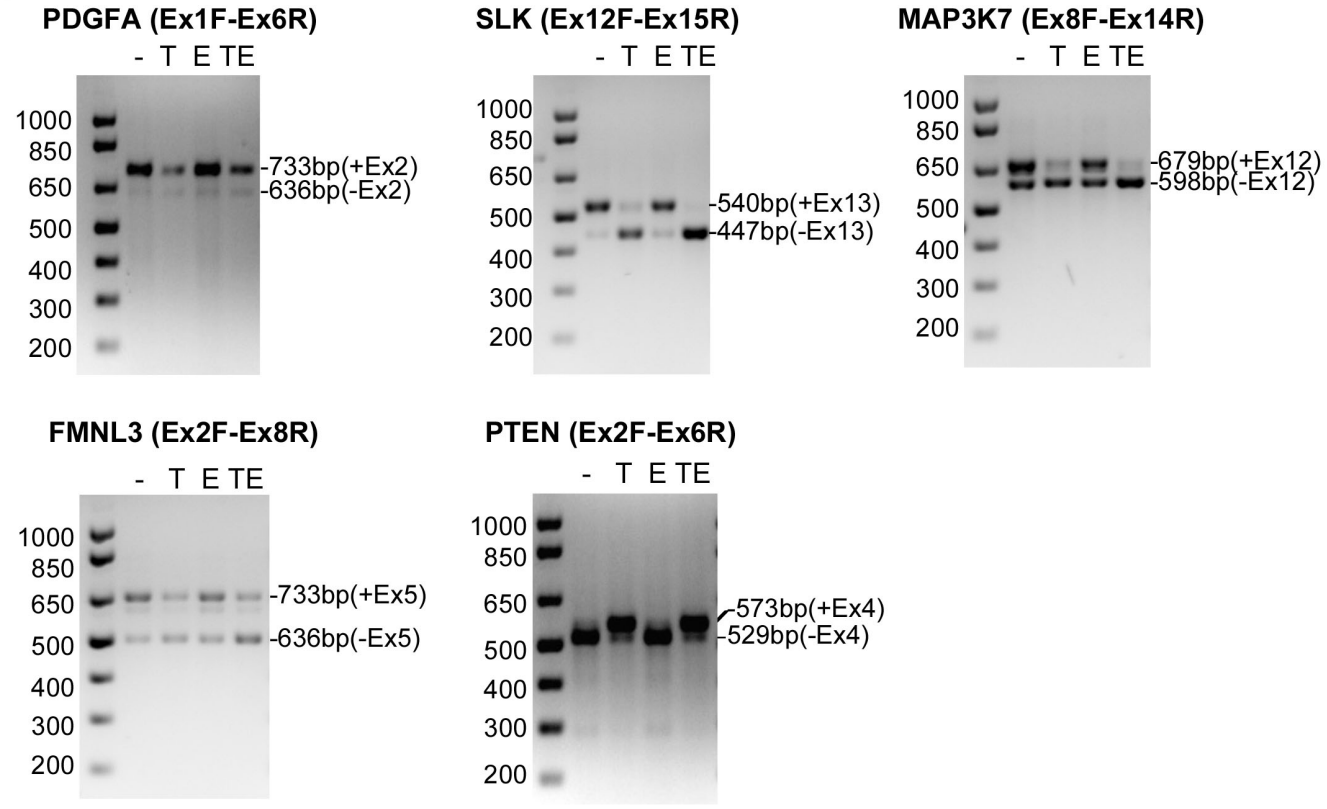


Supplemental Fig. S6, related to Fig. 6

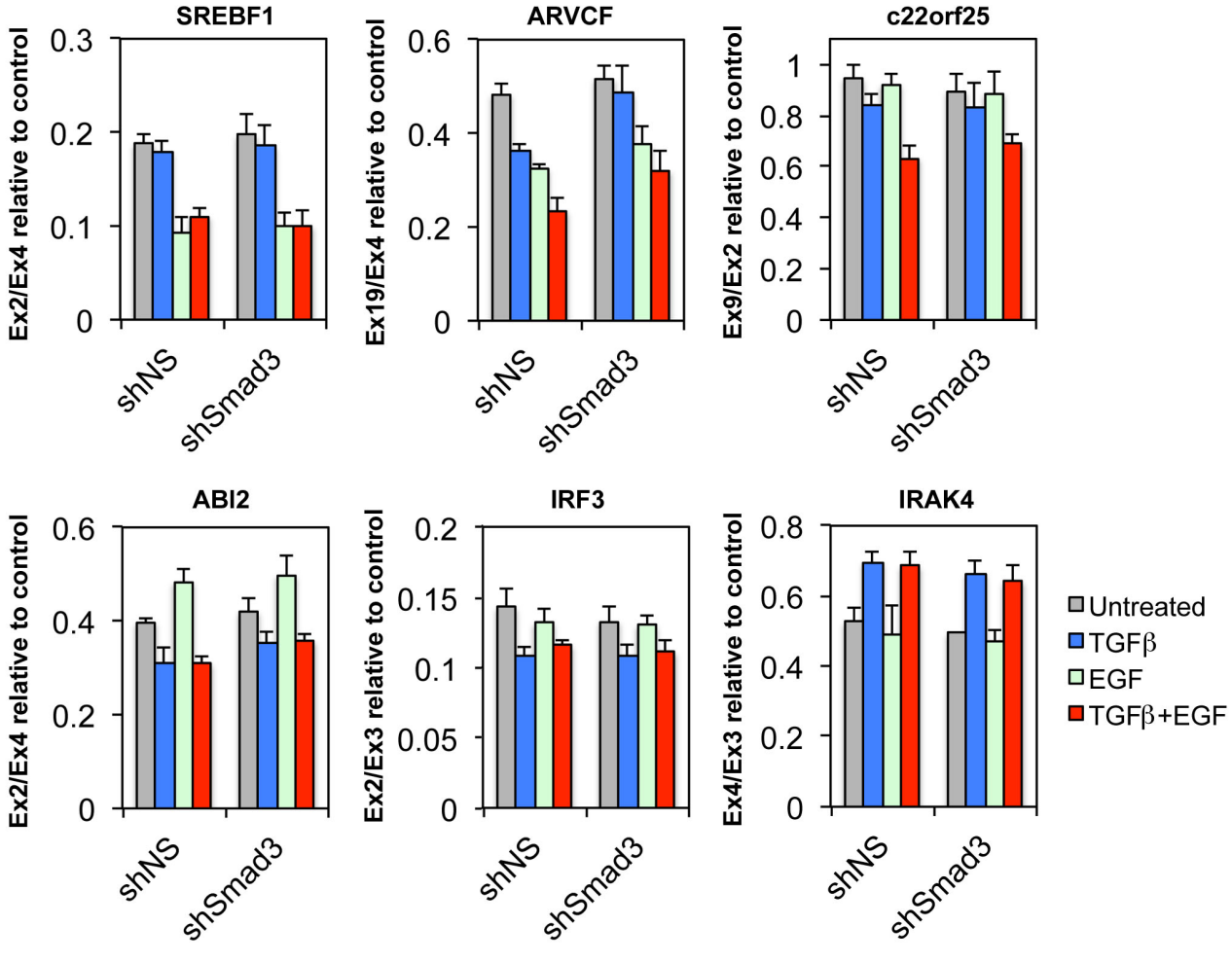


Supplemental Fig. S6, related to Fig. 7

A



B



Supplemental Table S1, related to experimental procedures and Fig. S3A: Primer sequences used.

Gene	Forward primer	Reverse Primer	Product size
CD44 Prmers			
Primers for qRT-PCR			
CD44 Total	CGCAGATCGATTTGAATATAACC	CCGATGCTCAGAGCTTTCTC	160
CD44E2	GACAAGTTTTGGTGGCACG	CACGTGGAATACACCTGCAA	105
CD44PreRNA	TCCCATCTTAGCCATTTAGGTC	GTTGTTTGCTGCACAGATGG	118
CD44Standard	AATCCCTGCTACCAGAGACC	TTCAGATCCATGAGTGGTATGGG	70
CD44e5-v8 Epithelial	TCCCTGCTACCAATATGGACTC	CAGAGTAGAAGTTGTTGGATGGTC	191
CD44v2	GCAACCAAGAGGCAAGAAA	CAGCCATTTGTGTTGTTGTGTG	96
CD44v3	CGTCTTCAAATACCATCTCAGC	CAATGCCTGATCCAGAAAAAC	93
CD44v4	TGACCACACAAAACAGAACC	GTTGTCTGAAGTAGCACTTCC	78
CD44v5	GAAATGGCACCACTGCTTATG	GTCTCTTCTTCCTCATGATGCT	88
CD44v6	AGGAACAGTGGTTTGGCAAC	CGAATGGGAGTCTTCTCTGG	68
CD44v7	TCAGCTCATACCAGCCATCC	TCCTTCTTCCTGCTTGATGAC	124
CD44v8	TCAGCCTACTGCAAATCCAA	GAGGTCCTGTCCTGTCCAAA	59
CD44v9	AGCAGAGTAATTCTCAGAGCTTC	TCAGAGTAGAAGTTGTTGGATGG	81
CD44v10	GGAATGATGTCACAGGTGGA	AGGTCACTGGGATGAAGGTC	123
Primers used in RNA-IP and ChIP			
Intron3-Exon4	TCCTACCTCATAGCTCCACCTG	TGGTAATTGGTCCATCAAAGG	70
Intron4-Exon5	TCCTTGTTTCCTGCCTCATC	GCCATCACGGTTAACAATAGC	225
Intron6-Exon7	GCTTTTTCCCTTTCTTTTCACC	TTGGTTGCTGTCTCAGTTGC	184
Intron7-Exon8	GCAAGTCAACTCCCTCACTTTC	TTGCATTAGTGGCTGAGCTG	117
Intron8-Exon9	CTTTCCCTCCATTTCCATGC	CCTGGTTCTGTTTTGTGTGG	192
Intron9-Exon10	CCAAGCCCATTGATTTTCTC	TGCTTCTGGGTTCCAGTTTC	236
Intron10-Exon11	ATCAGTGGCCTGTTTCCTTG	CATCTGTTGCCAAACCACTG	212
Intron11-Exon12	TCAAACTGCATGGTCACAGC	GTTGTCCTTCCTTGCATTGG	63
Intron12-Exon13	CACCTTCAGGTAGTGGTTTGG	CGCGTTGTCATTGAAAGAGG	160
Intron13-Exon14	TGTGTGCACAAACCTTGGTC	CTTCTTCCAAGCCTTCATGTG	186
Intron14-Exon15	TTCTGGAGACAAGCACATGG	TCCTTCGTGTGTGGGTAATG	173
Intron15-Exon16	CATTGTGGACCCATTCAAGG	TGGAATGTGTCTTGGTCTCC	180
Intron16-Exon17	TATGCAGCTCCACAAGGAAC	GTGTTTGCTCCACCTTCTTG	110
Intron17-Exon18	AAGTGTGTCTCTGAAGCTCACG	ATCAAAGCCAAGGCCAAGAG	92
Pol2 binding region in intron1	AAGTTTGTTGGGCAGGCT	GTTAGTCCAGACAACGACTGTG	113
Smad7 promoter	TAGAAACCCGATCTGTTGTTTGCG	CCTCTGCTCGGCTGGTTCCACTGC	133
Primers for DRB study			
Exon 2-3, intron3	CCTGCAGGTATGGGTTCATAGAAG	CACTTAAGCTGAGCTCCAAAGA	213
Exon3, intron3	TCCCATCTTAGCCATTTAGGTC	CACTTAAGCTGAGCTCCAAAGA	254
Exon11-12, intron12	CAACAGGGACAGCTGCAG	CAGCAGTCACTCAGTGGATTT	196
Exon12,intron12	TCAGCTCATACCAGCCATCC	CAGCAGTCACTCAGTGGATTT	186
Exon13-14,intron14	TCAATGACAACGCAGCAGAG	GATGCACCTCCACCAAAGA	193
Exon14,intron14	AGCAGAGTAATTCTCAGAGCTTC	GATGCACCTCCACCAAAGA	180
Exon16-17,intron17	GAATCAGATGGACACTCACATGG	TTCAAATGAGTCTCTTCGTCCTT	167
Exon17,intron17	CAAACACAACCTCTGGTCCTAT	TTCAAATGAGTCTCTTCGTCCTT	127
Primers to validate RNASeq results			
PDGFA			
Exon2	CGAGGTGATCGAGAGGCT	CGGAGTCTATCTCCAGGAGTC	77
Exon4	CAAGACCAGGACGGTCATTTA	GCAGCGTTTCACCTCCA	100
SLK			

Exon2	GGAGCCTTTGGGAAAGTGTA	TGGGTGATCACAAGATGCTAAT	141
Exon13	GCTGAAGAACCGAAAGAAGGA	TGAGCATGCTGGCTTTGT	114
MAP3K7			
Exon2	TGTTGGAAGAGGAGCCTTTG	ACGCTTTCCTCTCAGATTCAC	98
Exon12	CCTATTCCAAGCCTAAACGG	GATATGACGATCTCAGGGACA	79
FMNL3			
Exon2	CATGAACCTGCCTCCAGAC	GAATCGTTCCTGGTCACAGAT	88
Exon5	GTGGGTGCGGGAATTTCT	CCTCAAAGTCAAACATGACAGAAC	98
PTEN			
Exon2	GGATTCGACTTAGACTTGACCTATATT	CCTGTATACGCCTTCAAGTCTTT	84
Exon4	CTTTGTGCTGAAAGACATTATGACA	GTTCTAGCTGTGGTGGGTTATG	91
SREBF1			
Exon2	GAGCCATGGATTGCACTTTC	AGCATAGGGTGGGTCAAATAG	89
Exon4	AGCAGCTACTGACAGTCACA	CGGGACCTGCTGGATCT	80
ARVCF			
Exon4	ACCTGATGTGCTGGAGGA	GCCATCTTCGGATGTGACAATA	88
Exon19	GCCTCAGCCCGTTGATT	GATCCAAGCCCTAAGAACAAGA	77
c22orf25			
Exon2	CGGAGAAAGCCACAGAGAGAA	TTGATTTCAACAGCTGCCACTC	76
Exon9	GCCCATGCTGAGCAAGTA	CCGCATCTACCAGGATGATAGT	85
ABI2			
Exon2	GCTGTGTGGAGGATGGATTT	CATCACAACCATAGGGACGATG	81
Exon4	GGTGCGAGAAAGTGAGAGTAG	TGCAGTGGTATTGATCCTGTAG	91
IRF3			
Exon2	ACCTGGGACGCTGGAAA	CCACCTCCGCCTTCTCA	79
Exon3	TATGTTCCCGGGAGGGATAAG	GCTAAACGCAACCCTTCTTTG	89
IRAK4			
Exon3	GGAATAGAAGATGAACAAACCCATAAC	CTTCCATCCTTCTTGAGGATCAAT	109
Exon4	AACCATCTGGTGATGATAGATACAA	AAGATCCACAAGATCACCAACT	140
BAIAP2			
Exon6	AATACCAGACTGAGCAAAGGAG	TGTCCGAGTACTTCTGAGGAT	114
Exon14	ATTCTGAGCCGCCTGACTA	AGACCTGTCGTTGGTCACT	77
CLSTN1			
Exon4	CCCTTTGATGCAGTGGTAGTG	GGCCTGGATGGTGAATGAATAG	108
Exon11	TGAAACTGAGCCTGTGACTG	GAGAGTTAAGCCAGCCAGATT	88
KIF13A			
Exon6	GAGCAGCTGGGCCTTATT	CAACTTTAAAGGTCTGTGACTCATT	88
Exon38	CAGACCAGAAGCTGGAGAATC	GTTGATTCACCCAAGTGACAAAG	100
SCRIB			
Exon2	TTTCCGGCTGCTGAACT	CAGCTGCATGAAGTTGGC	88
Exon16	GTCGTTTGACCAGGCCAATAA	AGGATGGTGAGGGTCAGC	
Actin	AACTCCATCATGAAGTGTGACG	GATCCACATCTGCTGGAAGG	186
Primers for qRT-PCR			
P21	AGGCACTCAGAGGAGGTGAG	CCGCAGAAACACCTGTGAA	153
c-MYC	GGCTCCTGGCAAAAGGTCA	CTGCGTAGTTGTGCTGATGT	119
Vimentin	CAGGTGGACCAGCTAACCAA	TTCGGCTTCCTCTCTCTGAA	123
MMP2	GGAGCATGGCGATGGATAC	GTCCATAGCTCATCGTCATCAA	121
TWIST	GTCCGCAGTCTTACGAGGAG	GCTTGAGGGTCTGAATCTTGCT	156
Primers for RT-PCR on gel			
CD44	CAGCACTTCAGGAGGTTACAT	TTCAGATCCATGAGTGGTATGGG	559, 163

PDGFA	TGAATTTCGCCGCCACA	TGGTTGGCTGCTTTAGGT	733, 636
SLK	GCTTTGAGCAGGAAATGATGAG	GGCGCTTAAGTAGCTGATGT	540, 447
MAP3K7	GTACTCGACCACCACTGATAAA	GAGGTTGGTCCTGAGGTAGTA	679, 598
FMNL3	CATGAACCTGCCTCCAGAC	CTTGGATTCTTGTTATTGAGGCT	657, 573
PTEN	GGATTCGACTTAGACTTGACCTATATT	AGTTCCGCCACTGAACATT	573, 529
Sequences for ShRNA mediated knock down			
SMAD3			
Sense			
GATCCCCGGCCATCACCACGCAGAACTTCAAGAGAGTTCTGCGTGGTGATGGCCTTTTTA			
Antisense			
AGCTTAAAAAG GCCATCACCACGCAGAACTCTCTTGAAGTTCTGCGTGGTGATGGCCGGG			
PCBP1			
Sense			
GATCCCCGGGAGAGTCATGACCATTCTTCAAGAGAGAATG GTCATGACTCTCCCTTTTTA			
Antisense			
AGCTTAAAAAGGGAGAG TCATGACCATTCTCTCTTGAAGAATGGTCATGACTCTCCCGGG			
SMAD4			
Sense			
GATCCGTAGGACTGCACCATACACTTCAAGAGAGTGTATGGTGCAGTCCTACTTTTTTG			
Antisense			
TCGACAAAAAAGTAGGACTGCACCATACACTCTCTTGAAGTGTATGGTGCAGTCCTACG			
Primers Used for site directed mutagenesis			
PCBP1(S43A)	GATCCGCGCGCCAGCCTCCTCGCGGATC	GATCCGCGAGGAGGCTGGCGCGCGGGATC	
SMAD3 shRNA resistance	GGAGAAGGCCATCACAACTCAAAACGTCAA(	CTTGGTGTTGACGTTTTGAGTTGTGATGGCCTTCTCC	

Supplemental Table S2, related to Figure 7: RNA-seq data statistics.

	shNS			shNS(TE)*		
	replicate 1	replicate 2	replicate 3	replicate 1	replicate 2	replicate 3
Total_read_pairs	83877424	60874018	96436663	112675748	88531377	113428160
Read1_aligned	68576282	49834532	88072316	102935762	81406545	104538478
Read2_aligned	67735415	49229554	87866821	102661466	81154710	104249085
Overall_align%	81.30%	81.40%	91.20%	91.20%	91.80%	92.00%
Aligned_pairs	66217743	48115202	85792008	100301288	79226797	101837674
Aligned_pairs%	78.70%	78.70%	88.60%	88.70%	89.10%	89.40%
Read_pairs_junc	70093957	50948887	90147133	105295945	83334459	106949891
Read_align_junc%	83.57%	83.70%	93.48%	93.45%	94.13%	94.29%
PF_BASES	14792144981	10751026608	17966172395	21002393389	16578880734	21286831508
PF_ALIGNED_BASES	14791025199	10750239067	17964743499	21000734713	16577546075	21285303175
RIBOSOMAL_BASES	18690656	27365344	30917110	21522090	25279694	20647026
CODING_BASES	9500322381	6445967736	11634218014	13540024868	9699461721	13711875712
UTR_BASES	3939908108	2832650648	4800347121	5871660685	4618364780	5758331441
INTRONIC_BASES	610341578	574203752	651788478	744767460	983010189	787256249
INTERGENIC_BASES	721763729	870053724	847475193	822760980	1251431811	1007194047
PCT_RIBOSOMAL_BAS	0.13%	0.25%	0.25%	0.10%	0.15%	0.10%
PCT_CODING_BASES	64.23%	59.96%	64.76%	64.47%	58.51%	64.42%
PCT_UTR_BASES	26.64%	26.35%	26.72%	27.96%	27.86%	27.05%
PCT_INTRONIC_BASES	4.13%	5.34%	3.63%	3.55%	5.93%	3.70%
PCT_INTERGENIC_BAS	4.88%	8.09%	4.72%	3.92%	7.55%	4.73%
PCT_MRNA_BASES	90.87%	86.31%	91.48%	92.43%	86.37%	91.47%
PCT_USABLE_BASES	90.86%	86.30%		91.48%	92.43%	86.36%
	shPCBP1			shPCBP1(TE)		
	replicate 1	replicate 2	replicate 3	replicate 1	replicate 2	replicate 3
Total_read_pairs	101592511	95449576	92208906	98528226	86855899	105867575
Read1_aligned	93182310	87533759	84644898	89880634	75682234	91283505
Read2_aligned	92797354	87244100	84378635	89624802	75075839	90632295
Overall_align%	91.50%	91.60%	91.70%	91.10%	86.80%	85.90%
Aligned_pairs	90747793	85380784	82523131	87716856	72538817	87552882
Aligned_pairs%	88.90%	89.00%	89.10%	88.60%	83.10%	82.30%
Read_pairs_junc	95231873	89397077	86500405	91788581	78219258	94362920
Read_align_junc%	93.74%	93.66%	93.81%	93.16%	90.06%	89.13%
PF_BASES	18944717561	17799753483	17219201544	18291592880	15553430057	18766210060
PF_ALIGNED_BASES	18943150427	17798379512	17217860083	18290084168	15552254447	18764694186
RIBOSOMAL_BASES	34342424	69394878	73759290	61798870	31422716	38933480
CODING_BASES	12040257851	10963076902	10417821780	11225327767	9529274181	11546956140
UTR_BASES	5023253134	4569615723	4403309386	4711371767	4203157759	4976900856
INTRONIC_BASES	795621680	881700705	932078037	970275584	791758890	991559159
INTERGENIC_BASES	1049678069	1314596987	1390897540	1321315353	996643112	1210347456
PCT_RIBOSOMAL_BAS	0.18%	0.39%	0.43%	0.34%	0.20%	0.21%
PCT_CODING_BASES	63.56%	61.60%	60.51%	61.37%	61.27%	61.54%
PCT_UTR_BASES	26.52%	25.67%	25.57%	25.76%	27.03%	26.52%
PCT_INTRONIC_BASES	4.20%	4.95%	5.41%	5.30%	5.09%	5.28%
PCT_INTERGENIC_BAS	5.54%	7.39%	8.08%	7.22%	6.41%	6.45%
PCT_MRNA_BASES	90.08%	87.27%	86.08%	87.13%	88.30%	88.06%
PCT_USABLE_BASES	91.47%	90.07%	87.26%	86.07%	87.13%	88.05%
	shSMAD3			shSMAD3(TE)		
	replicate 1	replicate 2	replicate 3	replicate 1	replicate 2	replicate 3
Total_read_pairs	86486080	104216878	101938363	94761939	93393946	97722186
Read1_aligned	78661873	94965778	92588072	86157939	85352931	89626528
Read2_aligned	78359835	94661239	92186094	85851275	85077777	89332808
Overall_align%	90.80%	91.00%	90.60%	90.80%	91.20%	91.60%
Aligned_pairs	76671210	92478876	90063776	83975666	83269684	87436800
Aligned_pairs%	88.30%	88.30%	87.90%	88.30%	88.80%	89.10%
Read_pairs_junc	80350500	97148142	94710393	88033548	87161026	91522538
Read_align_junc%	92.91%	93.22%	92.91%	92.90%	93.33%	93.66%
PF_BASES	16022890278	19346214680	18861095520	17565643260	17370526006	18230852288
PF_ALIGNED_BASES	16021492514	19344817993	18859680462	17564319702	17369182336	18229515194
RIBOSOMAL_BASES	25539264	60547480	63094902	25682482	30763590	33209002
CODING_BASES	10406570946	12097332793	11705008356	11272208660	11091425825	11704221728
UTR_BASES	4226124561	4890036423	4800675872	4622083343	4548787801	4766267097
INTRONIC_BASES	559564383	917182072	905344654	747295925	796543708	765879109
INTERGENIC_BASES	803695494	1379723952	1385561809	897051135	901663744	959940674
PCT_RIBOSOMAL_BAS	0.16%	0.31%	0.33%	0.15%	0.18%	0.18%
PCT_CODING_BASES	64.95%	62.54%	62.06%	64.18%	63.86%	64.20%
PCT_UTR_BASES	26.38%	25.28%	25.45%	26.32%	26.19%	26.15%
PCT_INTRONIC_BASES	3.49%	4.74%	4.80%	4.25%	4.59%	4.20%
PCT_INTERGENIC_BAS	5.02%	7.13%	7.35%	5.11%	5.19%	5.27%
PCT_MRNA_BASES	91.33%	87.81%	87.52%	90.49%	90.05%	90.35%
PCT_USABLE_BASES	91.32%	87.81%	87.51%	90.49%	90.04%	90.34%

* {TE}: samples were treated with TGF-β and EGF.

Supplemental Table S3, related to Figue 7: Summary of Alternative Splicing Events.

Samples	Alternative Splicing Type	Total events	Significant* events
shPCBP1:	SE	69511	27
untreated vs TGFβ+EGF	RI	514	1
	MXE	20910	17
	A5SS	1492	4
	A3SS	1713	2
shNS vs shPCBP1	SE	67345	35
	RI	651	3
	MXE	19955	20
	A5SS	2185	11
	A3SS	2444	3
shNS: un vs T+E	SE	62034	239
untreated vs TGFβ+EGF	RI	486	8
	MXE	17697	97
	A5SS	1400	25
	A3SS	1637	15
shNS vs shSmad3	SE	66057	42
	RI	654	1
	MXE	19338	19
	A5SS	2178	8
	A3SS	2409	6
shSmad3:	SE	68791	54
untreated vs TGFβ+EGF	RI	515	1
	MXE	20590	31
	A5SS	1498	9
	A3SS	1715	6

*Changes with $|\Delta\psi|\geq 15\%$, Benjamini FDR ≤ 0.05 are considered as significant.

Supplemental Table S4, related to Figure 7:
Differential alternative splicing events between untreated and TGF-β and EGF treated HeLa cell lines.

Supplemental Table S5, related to Figure 7B:

Ingenuity Pathway enrichment analysis of genes whose alternative splicing changed ($|\Delta\psi| \geq 15\%$, FDR≤ 0.05) between untreated and TGF-β/EGF treated samples

Ingenuity Pathway Analysis (p≤0.02)

Ingenuity Canonical Pathways	-log(p-value)	Ratio	Overlaps with dataset	Molecules
RhoA Signaling	3.64E+00	6.56E-02	8/122 (7%)	ABL2,MYL6,PPP1R12A,BAIAP2,ARHGEF11,WASF1,SEPT6,ARHGAP8/PRR5-ARHGAP8
B Cell Receptor Signaling	3.22E+00	5.11E-02	9/176 (5%)	MAP3K9,PPP3CB,GAB1,NFATC3,VAV3,MAP3K7,PPP3CC,CREB3L4,PTEN
Netrin Signaling	2.76E+00	1.03E-01	4/39 (10%)	PPP3CB,NFATC3,PPP3CC,ENAH
Actin Cytoskeleton Signaling	2.59E+00	4.15E-02	9/217 (4%)	FN1,MYL6,PPP1R12A,PDGFA,VAV3,BAIAP2,WASF1,SSH1,IQGAP3
Epithelial Adherens Junction Signaling	2.47E+00	4.79E-02	7/146 (5%)	MYL6,FER,BAIAP2,CTNNA1,WASF1,PTEN,CTNND1
nNOS Signaling in Neurons	2.46E+00	8.51E-02	4/47 (9%)	PPP3CB,PPP3CC,CAPN7,PFKM
Inositol Pyrophosphates Biosynthesis	2.45E+00	2.86E-01	2/7 (29%)	PPIP5K1,IP6K2
Reelin Signaling in Neurons	2.39E+00	6.33E-02	5/79 (6%)	MAP3K9,LRP8,ARHGEF11,ARHGEF10,APP
Phospholipase C Signaling	2.32E+00	3.77E-02	9/239 (4%)	PPP3CB,MYL6,PPP1R12A,NFATC3,HDAC10,ARHGEF11,PPP3CC,CREB3L4,ARHGEF10
PKCÎ¸ Signaling in T Lymphocytes	2.30E+00	5.08E-02	6/118 (5%)	MAP3K9,PPP3CB,NFATC3,VAV3,MAP3K7,PPP3CC
CD27 Signaling in Lymphocytes	2.30E+00	7.69E-02	4/52 (8%)	MAP3K9,APAF1,MAP3K7,CASP8
HIPPO signaling	2.24E+00	5.81E-02	5/86 (6%)	PPP1R12A,CD44,SCRIB,TEAD2,RASSF1
SAPK/JNK Signaling	2.08E+00	5.32E-02	5/94 (5%)	SH2D2A,MAP3K9,GAB1,NFATC3,MAP3K7
RhoGDI Signaling	2.07E+00	4.05E-02	7/173 (4%)	MYL6,PPP1R12A,CD44,ARHGEF11,WASF1,ARHGEF10,ARHGAP8/PRR5-ARHGAP8
Signaling by Rho Family GTPases	1.88E+00	3.42E-02	8/234 (3%)	MAP3K9,MYL6,PPP1R12A,BAIAP2,ARHGEF11,WASF1,SEPT6,ARHGEF10
Methylglyoxal Degradation VI	1.88E+00	1.00E+00	1/1 (100%)	LDHD
Sulfite Oxidation IV	1.88E+00	1.00E+00	1/1 (100%)	SUOX
Protein Kinase A Signaling	1.84E+00	2.86E-02	11/385 (3%)	FLNB,AKAP13,HHAT,PPP3CB,MYL6,PPP1R12A,NFATC3,AKAP9,PPP3CC,CREB3L4,PTEN
DNA Double-Strand Break Repair by Non-Homologous End Joining	1.84E+00	1.43E-01	2/14 (14%)	DCLRE1C,XRCC4
Role of PKR in Interferon Induction and Antiviral Response	1.80E+00	7.50E-02	3/40 (8%)	APAF1,MAP3K7,CASP8
Leukocyte Extravasation Signaling	1.77E+00	3.54E-02	7/198 (4%)	MYL6,VAV3,FER,CD44,CTNNA1,ARHGAP8/PRR5-ARHGAP8,CTNND1
Toll-like Receptor Signaling	1.77E+00	5.41E-02	4/74 (5%)	SIGIRR,MAP3K7,TIRAP,IRAK4
Granzyme B Signaling	1.73E+00	1.25E-01	2/16 (13%)	APAF1,CASP8

Supplemental Experimental Procedures

Cell lines, culture conditions and chemical inhibitors

HeLa and OSCC3 cells were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS) and 100 unit/ml penicillin-streptomycin. MCF10A cells were maintained in DMEM/F12 medium supplemented with 5% horse serum, 10 μg/ml insulin, 20 ng/ml EGF, 500 ng/ml hydrocortisone, and 100 ng/ml cholera toxin. HMLE cells were obtained from Dr. Jing Yang (University of California, San Diego) and cultured as described (Elenbaas et al., 2001). For treatment ≤ 2hr, cells were stimulated with 4 ng/ml TGF-β or 10 nM EGF or both after culturing under serum-starved condition in DMEM supplemented with 0.2% FBS for 4 hr. TGF-α was used at final concentration of 10 ng/ml. For long-term treatment, cells were not serum starved. MEK1/2 inhibitor U0126 and AKT inhibitor VIII were obtained from EMD Millipore, CDK inhibitor flavopiridol and EGFR inhibitor Afatinib were from Santa Cruz Biotech.

Plasmid Constructions and transfections

To knock down SMAD3, SMAD4 and PCBP1 expression, previously validated shRNA hairpin sequences were inserted in pRetro-HIG, pSuper-retro-puro, or pSuper retro-neo-GFP vectors, and transfected in HeLa cells and selected depending on the selection requirement. pSuper retro-neo-GFP was used as the vector control. To generate shSMAD3-resistant pools of HeLa cells, SMAD3 WT and SMAD3 T179V cDNAs (Tang et al., 2011) in the pLPCX vector were modified with silent mutations in codons corresponding to the shSMAD3 target sequence with Site Directed Mutagenesis Kit (Agilent Technologies). T7-PCBP1 and T7-PCBP1-T60A-T127A (Mut) plasmids were generously gifts from Dr. Kumar (George Washington University). The T7-PCBP1 S43A mutant was generated by site directed mutagenesis. Plasmids pRRL-shCD44-2 and pBabe-puro-CD44s used for knocking down total CD44 or expressing CD44s, respectively, were obtained from Addgene. pBabe-CD44v is a gift from Dr. Chonghui Cheng (Brown et al., 2011). Sequences of primers used in the construction of shRNA vectors and site directed mutagenesis are listed in Supplementary Table S1. Plasmid transfections were carried out using Lipofectamine 2000 (Invitrogen) reagent according to manufacturer's instructions.

Mass spectroscopic identification of pT179 peptide binding proteins

Cells were first treated with 10 μM TβRI inhibitor SB431542 (Tocris Biosciences) in DMEM supplemented with 0.2% FBS for 4 hr, then washed with PBS twice before lysed by incubating on ice for 30 min in buffer 1 (10 mM Tris-HCl, pH 7.4, 10 mM KCl, 0.4% NP40, 0.1 mM EDTA, supplemented with 20 mM n-ethylmaleimide). Nuclei were pelleted by 30 min centrifugation at 1000xg and washed twice in buffer 1, then resuspended in buffer 2 (50 mM Tris-Hcl, pH 7.4, 600 mM KCl, 20% glycerol, 0.1 mM EDTA, supplemented with 20 mM n-ethylmaleimide) with 30 min incubation on ice followed by sonication. Nuclear lysates were cleared by centrifugation at 20,000xg for 30 min, and the supernatant was combined at 1:1 ratio with buffer 3 (50 mM Tris-HCl, pH 7.4, 0.1 mM EDTA, 20 mM n-ethylmaleimide). Affinity resins conjugated with SMAD3 T179 or pT179 peptides were described previously (Tang et al., 2011). Immobilized peptides were incubated with nuclear lysates overnight and washed with buffer 4 (50 mM Tris-HCl, pH 7.4, 300 mM KCl, 10% glycerol, 0.1 mM EDTA). Proteins bound to the peptides were resolved on SDS-PAGE and stained with PageBlue (Fermentas). Gel bands that were only visible in the phosphorylated pT179 peptide lanes were excised and subjected to trypsin digestion. The tryptic peptides were extracted and analyzed by a coupled liquid chromatography-mass spectrometer.

Invasion assay

Cell invasion assay across a basement membrane was performed using Corning Biocoat™ Growth Factor Reduced Matrigel Invasion Chamber with 8.0 mm PET membrane in 24-well plates (Corning Life Sciences). Briefly, a volume of 2.5×10^4 cells in DMEM medium that had been treated with indicated growth factors for 24 hr was added to each invasion chamber. After 18 hr, invaded cells located on the underside of the chamber were fixed with formaldehyde for 10 min and stained with trypan blue before counting.

Nuclear and cytosolic fractionation

Cells grown in 10 cm plates were lysed with buffer A (10 mM HEPES. pH 7.5, 1.5 mM MgCl2, 10 mM KCl, 5 mM EDTA, 0.5 mM DTT, 0.2% NP-40) for 15 min at 4°C on a rocking shaker after washed twice with ice cold PBS without calcium and magnesium. The lysates were gently collected from the plate and cleared by centrifugation in Eppendorf tubes at 2000g for 5 min. The resulting supernatants (cytosolic fraction) were transferred to new tubes. After gently washing the rest of the plate with buffer A, 350 ul RIPA buffer (50 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.25% Sodium deoxycholate, 1% NP-40, and 1 mM EDTA) were added and incubated at 4°C for 30 min

before centrifugation at 12,000g for 5 min. The supernatant from the above were then transferred to new tubes (nuclear fraction). All buffers were supplemented with protease cocktail and phosphatase inhibitors.

Western blot and immunoprecipitation

Cells were lysed in NP-40 lysis buffer (50 mM Tris-HCl, pH 7.4, 150 mM NaCl, 1% NP-40) supplemented with Protease Inhibitor Cocktail (Roche) prior to centrifugation at 4°C for 20 min at 12,000g. Cell lysates were quantitated by BCA protein assay before being subjected to SDS-PAGE and Western blot. Antibodies used are SMAD3 (Abcam), SMAD3 pS423/S425 (Rockland), SMAD3 pT179 (Rockland), PCBP1 (MBL international), FLAG M2 (Sigma), U2AF2 (MBL international), PRP6 (Abcam), pAKT, GSK3beta, GAPDH (Santa Cruz Biotechnology), HSC70 and Lamin B (Santa Cruz Biotechnology). For immunoprecipitation, antibodies were added into precleared cell lysates and incubated for 4 hr at 4°C before the protein G-Sepharose slurry (Life Technologies) was added and further incubated overnight. Beads were then washed 4-5 times with NP-40 buffer and subjected to SDS-PAGE and Western blot.

Immunofluorescence imaging

Cells grown in chamber slides were washed with PBS, fixed with 4.0% paraformaldehyde for 20 min, then permeabilized with 0.2% Triton-X for 5 min. Slides were incubated with mouse anti-SMAD3, rabbit anti-PCBP1, and mouse anti-SC35 (Sigma) primary antibodies for overnight at 4°C. Secondary antibodies used were anti-mouse Alexa Fluor 488, anti-goat Alexa Fluor 568, and anti- mouse Alexa Fluor 650 (Molecular probe). Stained Cells were examined with a Zeiss LSM 780 Rm K confocal system (Carl Zeiss Inc) with an Axiovert 200 M inverted microscope. Images were collected in Zeiss AIM software (v. 4.0) using a 63x Plan-Apochromat 1.4 NA oil immersion objective lens. Images were 512 x 512 pixels with line averaging of 4.

Proximity Ligation Assay

Proximity ligation assay was performed to visualize the interaction between PCBP1 and SMAD3 using a Duolink In Situ kit (Olink Biosciences). Briefly, fixed and permeabilize cells were blocked with blocking solution for 30 min at 37°C followed by incubation with mouse anti-SMAD3 and rabbit anti-PCBP1 primary antibodies for 1 hr in a humidified chamber. After washing, secondary antibodies conjugated with PLA probe PLUS (for mouse antibodies) and PLA probe MINUS (for rabbit antibodies) were added and incubated in a pre-heated humidifier chamber for 1 hr at 37°C. Ligation solution containing two oligonucleotides specific for PLUS and MINUS probes and ligase were added. Oligonucleotides hybridized to PLA probes and joined to make a close circle if they were in close proximity. A solution containing fluorescent oligonucleotides and polymerase was further added to amplify the signal for detection by microscopy.

Quantitative RT-PCR

RNA was isolated with RNeasy mini kit (Qiagen). Precipitated RNA was resuspended in nuclease-free water and quantified using a Nanodrop. High Capacity Reverse Transcription Kit (ABI, Life Tech) was used to generate cDNA from RNA (500 -2000 ng). qRT-PCR was performed with Power SYBR Green PCR Master Mix (Life Technologies) using specific oligonucleotide primers as specified (Table S1). All samples were run in three biological replicates, each with two-three technical replicates. Fold-changes were calculated using the 2^-ddCt method, and the calculated threshold values were determined by the maximum curvature and ΔCt was calculated as $Ct^{Control} - Ct^{sample}$. All qRT-PCR values were normalized against β-actin.

RNA immunoprecipitation (RIP)

RIP was done using RIP-Assay Kit (MBL International). Briefly, precleared cell lysate was incubated with RIP certified antibodies immobilized on magnetic Protein G Dynabeads (Life Technologies) for 4 hr at 4°C. The bound RNA was recovered and quantified using qRT-PCR. To determine differential binding between the variable and standard region, RNACHIP-IT kit (Active Motif) was used, and cells were cross-linked with UV crosslinking with 200 mJ/cm^2 at 254 nm using CL-1000 ultraviolet crosslinker (UVP Inc.). Nuclear fraction was prepared and RNA was fragmented with Biorupture (Diagenode) for 20 min (30sec on/off), which gives rise to RNA fragments of 50-800 bases in length. These fragmented RNAs were used for immunoprecipitation. After precipitation, beads were washed and Protein-RNA complex was eluted and treated with Proteinase K for 1 hr at 42°C. RNA was then isolated, incubated with DNase I to eliminate any genomic DNA contamination and subjected for qRT-PCR analysis.

Chromatin immunoprecipitation (ChIP)

ChIP assays were carried out with an EZ-ChIP Chromatin Immunoprecipitation Kit (Millipore). Sonication was performed using QSonica Q500 instrument and a microtip probe with 10 sec pulse and 50 sec rest in between for 4-5 cycles with output intensity set at 30 %. One-tenth of each chromatin sample was kept as the input DNA control. Immunoprecipitations were carried out using anti-RNAPII, anti-SMAD3 and an isotype-matched IgG as the control.

Inhibition and re-initiation of transcription

DRB (5,6-dichlorobenzimidazole 1β-d-ribofuranoside, Sigma) was used to reversibly block transcription in cells as described (Singh and Padgett, 2009). TGF-β and/or EGF were added 30 min before the end of DRB treatment. Following the incubation with 100 μM DRB for 4 hr, cell were washed twice with PBS, incubated in fresh medium with indicated factors for various time periods, and then harvested for the isolation of total RNAs.

Luciferase reporter assays

Reporter assays were performed in 12 well plates using CAGA12-Luc (0.5 μg) and pRL-TK (0.2 μg) reporter plasmids, and the luciferase activities were determined using Dual Luciferase Reporter Assay System (Promega) as described (Tang et al., 2011).

FACS analysis

After washed with PBS, cells were resuspended in ice cold FACS Buffer (PBS, 0.5% BSA 0.1% NaN3) at $1x10^6$ cells/ml. 0.1 ml cell suspension was transferred to polystyrene round-bottom tubes, and 0.1 ml blocking buffer (3.0% BSA in PBS) was added to each sample. After incubation on ice for 20 min, the cells were collected by centrifugation at 1500 rpm for 5 min at 4°C. 1-10 μg/ml of PE-CD44 total (BD biosciences) or FITC-CD44std (ebiosciences) primary labeled antibodies were added in diluted FACS buffer, then incubated for at least 30 min at 4°C in the dark followed by washing 3 times. The cells were then resuspended in 500 μl of ice-cold FACS buffer for analysis. For primary unlabeled CD44 epithelial antibody (Cosmo Bio USA), the FITC labeled secondary antibody was added to cells in FACS buffer, incubated for 30 min at 4°C in the dark, then the cells were resuspended in 500 μl of ice-cold FACS buffer for analysis. All isotype controls were from BD biosciences.

Tumor xenograft and metastasis assay

For tumor xenograft, 5-6 weeks old female athymic nu/nu mice were inoculated subcutaneously on both hinder flank with 0.1 ml of HeLa cells ($1x10^6$) per flank. Tumor volume were measured with calipers and calculated by the formula ($0.5*L*S^2$), where L and S were the long and short dimensions, respectively. For metastasis assay, HeLa cells ($2x10^6$) were injected intravenously through the tail vein of 5-6 weeks old female athymic nu/nu mice. Three weeks after injection, mice were euthanized and examined grossly at necropsy for the presence of metastases in internal organs. Whole lung from each animal were inflated and fixed in 10% neutral buffered formalin before paraffin enbedding. Quantification of metastases were performed using H&E stained serial sections. All mice were maintained and handled according to protocols approved by the Animal Care and Use Committee of National Cancer Institute.

RNA-Seq analyses

RNAs from Hela cells stably transfected with shNS, shSMAD3 or shPCBP1 were sequenced on HiSeq2000 with 2 samples pooled in to one lane with Illumina TruSeq v3 chemistry. Demultiplexing was done allowing 1 mismatch in the barcodes. Paired-ended RNA-Seq sequencing reads were aligned with TopHat version 2.0.9 (Kim et al., 2013; Trapnell et al., 2009) (http://tophat.cbcb.umd.edu/) against the reference human genome hg19, with UCSC known gene transcripts as the gene model annotation. Quality of the RNA-seq data was assessed using Picard's CollectRnaSeqMetrics package on alignment bam files (http://broadinstitute.github.io/picard/). Mapping statistics for the RNA-Seq data were listed in Supplemental Table S2. In order to identify differential alternative splicing events between untreated and treated cell lines, rMATS tool version 3.0.8 (Shen et al., 2014) (http://rnaseq-mats.sourceforge.net/) was used with TopHat generated alignment bam files as input in default options, and UCSC known gene transcripts as the gene annotation model. Splicing events were evaluated using sequencing reads that span the splicing junctions and on target exons. The differential splicing events in rMATS output files were further filtered using the following criteria in order to identify significant events: the inclusion and exclusion reads counts are greater than or equal to 10 in any of the comparing samples; the threshold of exon inclusion levels between two treatments, $|\Delta\psi| \geq 15\%$, and false discovery rate (FDR) of $\leq 5\%$. RNA-seq data sets were deposited in GEO with accession number GSE71419.

Statistical analysis

Two-tailed student's t-test was used for statistical analysis. All bar graphs are displayed as mean ± S.D.

Supplemental Reference

Elenbaas, B., Spirio, L., Koerner, F., Fleming, M.D., Zimonjic, D.B., Donaher, J.L., Popescu, N.C., Hahn, W.C., and Weinberg, R.A. (2001). Human breast cancer cells generated by oncogenic transformation of primary mammary epithelial cells. Genes Dev *15*, 50-65.

Kim, D., Pertea, G., Trapnell, C., Pimentel, H., Kelley, R., and Salzberg, S.L. (2013). TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions. Genome biology *14*, R36.

Shen, S., Park, J.W., Lu, Z.X., Lin, L., Henry, M.D., Wu, Y.N., Zhou, Q., and Xing, Y. (2014). rMATS: robust and flexible detection of differential alternative splicing from replicate RNA-Seq data. Proc Natl Acad Sci U S A *111*, E5593-5601.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: discovering splice junctions with RNA-Seq. Bioinformatics *25*, 1105-1111.