

# **TGF- $\beta$ -SPTBN1-CTCF-regulated tumor suppression in Beckwith-Wiedemann syndrome, a human stem cell disorder**

Jian Chen, Zhi-Xing Yao, Jiun-Sheng Chen, Young Jin Gi, Nina M. Muñoz, Suchin Kundra, H. Franklin Herlong, Yun Seong Jeong, Alexei Goltsov, Kazufumi Ohshiro, Nipun A. Mistry, Jianping Zhang, Xiaoping Su, Sanaa Choufani, Abhisek Mitra, Shulin Li, Bibhuti Mishra, Jon White, Asif Rashid, Alan Yaoqi Wang, Milind Javle, Marta Davila, Peter Michaely, Rosanna Weksberg, Wayne L. Hofstetter, Milton J Finegold, Jerry W. Shay, Keigo Machida, Hidekazu Tsukamoto, and Lopa Mishra

Correspondence to:

Lopa Mishra, MD  
Director, Center for Translational Research  
Department of Surgery and GW Cancer Center,  
George Washington University, 2150 Pennsylvania  
Avenue, NW, Washington, DC 20037  
& The Del & Dennis McCarthy Distinguished Professor  
in Gastrointestinal Cancer Research, Department of  
Gastroenterology, Hepatology, and Nutrition – Unit 1466  
The University of Texas MD Anderson Cancer Center  
1515 Holcombe Boulevard, Houston, TX 77030  
lmishra@mdanderson.org, lmishra@GWU.edu  
Tel: 240-401-2916, 202-741-3225

## **Supplemental Materials and Methods**

### **Whole-transcriptome Sequencing Analysis.**

Whole-transcriptome RNA sequencing were performed and analyzed at the MD Anderson Cancer Center DNA core facility. We generated a standard bam file by aligning the raw paired-end reads in FASTQ format to the human reference genome, GRCh37/hg19, using MOSAIK alignment software (1). We then counted the uniquely mapped reads in genomic features such as genes annotated in refSeq to generate the raw counts for each gene, where raw reads count data were subsequently normalized across samples using DESeq with variance stabilization (2).

### **Database analyses.**

Affymetrix mRNA microarray data GSE3228 were downloaded from NCBI's Gene Expression Omnibus (GEO). *Smad3*<sup>-/-</sup> primary keratinocytes were derived from *Smad3* null mice. The transcriptome data of two independent *Smad3*<sup>-/-</sup> keratinocytes were downloaded and further normalized with wild type mouse keratinocytes. Most of the decreased TGF- $\beta$ /Smad3-regulated genes presented on Figure 2A are validated by IPA analysis (cutoff: adjusted *P* value < 0.05) in *Smad3*<sup>-/-</sup> primary keratinocytes. The information about somatic mutations in *SPTBN1* and *SMAD3* reported in the Catalogue of Somatic Mutations in Cancer and The Cancer Genome Atlas (TCGA) liver cancers was downloaded in May 2014 (Supplemental Table 3).

### **Preparation of Nuclear and Cytoplasmic Proteins.**

Nuclear and cytoplasmic proteins were prepared as previously described (3). Briefly, the cells were harvested and incubated in buffer A (10 mmol/L Hepes, pH 7.8, 10 mmol/L KCl, 0.1 mmol/L EDTA, 1 mmol/L dithiothreitol, 2 mg/ml aprotinin, 0.5 mmol/L phenylmethyl sulfonyl fluoride, and 0.5% Triton X-100). After centrifugation at 5000 rpm, buffer A was collected,

including cytoplasmic proteins, as the cytoplasmic protein sample. Buffer C (50 mmol/L Hepes, pH 7.8, 420 mmol/L KCl, 0.1 mmol/L EDTA, 5 mmol/L MgCl<sub>2</sub>, 10% glycerol, 1 mmol/L dithiothreitol, 2 mg/ml aprotinin, and 0.5 mmol/L phenylmethyl sulfonyl fluoride) was added to the pellet. After rotation for 30 minutes and centrifugation at 15,000 rpm, supernatants were collected as the nuclear protein sample. Then, lysates were immunoprecipitated with the indicated antibodies. The following antibodies were used for immunoblotting and immunoprecipitation analyses: V5 antibody (R961-25, Invitrogen), Flag M2 antibody (Sigma Aldrich, F3165), His antibody (2366, Cell Signaling), Actin antibody (A2066, Sigma Aldrich), Tubulin antibody (T8328, Sigma Aldrich), c-Myc antibody (sc-40, Santa Cruz Biotechnology), CTCF antibody (2899, Cell Signaling), Smad3 antibody (9523, Cell Signaling), phosphor-Smad3 antibody (Ser423/425, 9520, Cell Signaling), Smad2 antibody (3122, Cell Signaling), phosphor-Smad2 antibody (Ser465/467, 8828, Cell Signaling),  $\beta$ 2SP antibody (customized antibody from bioSythesis).

#### **Primers for Quantitative RT-PCR Analyses and ChIP Assays.**

The sequences used for quantitative RT-PCR analysis were as follows:

*hTERT*, forward 5'- AAATGCGGCCCTGTTTCT-3'; reverse 5'- CAGTGCGTCTTGAGGAGCA-3'. *mTERT*, forward 5'- TCTACCGCACTTTGGTTGCC-3'; reverse 5'- CAGCACGTTTCTCTCGTTGC-3'. *IGF2*, forward 5'- GTGGCATCGTTGAGGAGTG-3'; reverse 5'- CACGTCCCTCTCGGACTTG-3'. *CTCF*, forward 5'- CAGTGGAGAATTGGTTCGGCAT-3'; reverse 5'- CTGGCGTAATCGCACATGGA-3'. *Smad3*, forward 5'- TGGACGCAGGTTCTCCAAAC-3'; reverse 5'- CCGGCTCGCAGTAGGTAAC-3'. *m HEY1*, forward 5'- CCGACGAGACCGAATCAATAAC-3'; reverse 5'- TCAGGTGATCCACAGTCATCTG-3'.

*mESR1*, forward 5'- CCCGCCTTCTACAGGTCTAAT-3'; reverse 5'-  
CTTTCTCGT TACTGCTGGACAG-3'. *mRunx3*, forward 5'-  
GACTCCTTCCCCAACTATACACC-3'; reverse 5'- CGCTGTTCTCGCCCATCTT-3'.  
*mSLC1A2*, forward 5'- GCCAACAATATGCCCAAGCAG-3'; reverse 5'-  
GACACCAAACACAGTCAGTGA-3'. *hRUNX1*, forward 5'- CTGCCCATCGCTTTCAAGGT-  
3'; reverse 5'- GCCGAGTAGTTTTTCATCATTGCC-3'. *hCDKN1A*, forward 5'-  
TGTCCGTCAGAACCCATGC-3'; reverse 5'-AAAGTCGAAGTTCCATCGCTC-3'.  
*hCDKN1B*, forward 5'-AACGTGCGAGTGTCTAACGG-3'; reverse 5'-  
CCCTCTAGGGGTTTGTGATTCT-3'. *hSERPINE1*, forward 5'-  
ACCGCAACGTGGTTTTTCTCA-3'; reverse 5'- TTGAATCCCATAGCTGCTTGAAT-3'.  
*hIRF7*, forward 5'- GCTGGACGTGACCATCATGTA-3'; reverse 5'-  
GGGCCGTATAGGAACGTGC-3'. *hFOS*, forward 5'-CCGGGGATAGCCTCTCTTACT-3';  
reverse 5'-CCAGGTCCGTGCAGAAGTC-3'. *hTGIF1*, forward 5'-  
GGGATTGGCTGTATGAGCACC-3'; reverse 5'- GGCGGGAAATTGTGAACTGA-3'.  
*hMMP9*, forward 5'- TGTACCGCTATGGTTACTCTCG-3'; reverse 5'-  
GGCAGGGACAGTTGCTTCT-3'.  
*hMYC*, forward 5'-GGCTCCTGGCAAAAGGTCA -3'; reverse 5'-  
CTGCGTAGTTGTGCTGATGT-3'.  
*Gapdh*, forward 5'-TGCACCACCAACTGCTTAGC-3';  
*Gapdh*, reverse 5'-GTCTTCTGGGTGGCAGTGATG-3'.  
*hALB*, forward 5'-TGCAACTCTTCGTGAAACCTATG-3';  
*hALB*, reverse 5'-ACATCAACCTCTGGTCTCACC-3';  
*mALB*, forward 5'-CAAGAGTGAGATCGCCCATCG-3';

*mALB*, reverse 5'-TTACTTCCTGCACTAATTTGGCA-3';  
*hKRT19*, forward 5'-AACGGCGAGCTAGAGGTGA-3';  
*hKRT19*, reverse 5'-GGATGGTCGTGTAGTAGTGGC-3';  
*mKRT19*, forward 5'-GTTCAGTACGCATTGGGTCAG-3';  
*mKRT19*, reverse 5'-GAGGACGAGGTCACGAAGC-3';  
*hPROM1(CD133)*, forward 5'-AGTCGGAAACTGGCAGATAGC-3';  
*hPROM1(CD133)*, reverse 5'-GGTAGTGTTGTACTGGGCCAAT-3';  
*mPROM1(CD133)*, forward 5'-ACTGGGGCTGTGTGGAAAG-3';  
*mPROM1(CD133)*, reverse 5'-GCATTGAAGGTATCTTGGGTCTC-3';  
*hNANOG*, forward 5'-TTTGTGGGCCTGAAGAAAAC-3';  
*hNANOG*, reverse 5'-AGGGCTGTCCTGAATAAGCAG-3';  
*mNANOG*, forward 5'-CACAGTTTGCCTAGTTCTGAGG-3';  
*mNANOG*, reverse 5'-GCAAGAATAGTTCTCGGGATGAA-3';  
*hPOU5F1(OCT4)*, forward 5'-CTGGGTTGATCCTCGGACCT-3';  
*hPOU5F1(OCT4)*, reverse 5'-CCATCGGAGTTGCTCTCCA-3';  
*mPOU5F1(OCT4)*, forward 5'-AGAGGATCACCTTGGGGTACA-3';  
*mPOU5F1(OCT4)*, reverse 5'-CGAAGCGACAGATGGTGGTC-3';  
*hSOX2*, forward 5'-GCCGAGTGGAACTTTTGTCG-3';  
*hSOX2*, reverse 5'-GGCAGCGTGTACTTATCCTTCT-3';  
*mSOX2*, forward 5'-GCGGAGTGGAACTTTTGTC-3';  
*mSOX2*, reverse 5'-GGGAAGCGTGTACTTATCCTTCT-3';  
*hEPCAM*, forward 5'-AATCGTCAATGCCAGTGTACTT-3';  
*hEPCAM*, reverse 5'-TCTCATCGCAGTCAGGATCATAA-3';

*mEPCAM*, forward 5'-CTGGCGTCTAAATGCTTGGC-3';

*mEPCAM*, reverse 5'-CCTTGTCGGTTCTTCGGACTC-3';

*hAFP*, forward 5'-CTTTGGGCTGCTCGCTATGA-3';

*hAFP*, reverse 5'-GCATGTTGATTTAACAAGCTGCT-3';

*mAFP*, forward 5'-AGCTTCCACGTTAGATTCCTCC-3';

*mAFP*, reverse 5'-ACAAACTGGGTAAAGGTGATGG-3';

The sequences used for ChIP assays were as follows (Supplementary Figure S10):

hTERT ChIP-1-F: CTCTGCAGTCCGAGGCTTG; hTERT ChIP-1-R:

CCCGTCATTTCTCTTTGCAG. hTERT ChIP-2-F: CCGGACCTGGAGGCAGCC; hTERT

ChIP-2-R: CCCAGCGGAGAGAGGTC. hTERT ChIP-3-F: ACCTCTCTCCGCTGGGGC;

hTERT ChIP-3-R: CGTCTGTGCCCGCGAATC. hTERT ChIP-4-F:

GGACCGCGCTTCCCACGT; hTERT ChIP-4-R: AGGAAGGGGAGGGGCTGG. hTERT

ChIP-5-F: AGCCCCTCCCCTTCCTTTCC;

hTERT ChIP-5-R: AGCGCACGGCTCGGCAGC. hTERT ChIP-6-F:

AGCCGTGCGCTCCCTGCT; hTERT ChIP-6-R: GCACGCACACCAGGCACT. mTERT

ChIP-1-F: AAATTCAGCAGCCCCTCTCT; mTERT ChIP-1-R:

TCTCAACTCACGCACCCATA. mTERT ChIP-2-F: ATGGTCGCACCACAATAAAG;

mTERT ChIP-2-R: GGGCAAGCAAAGAAGCCTAT. mTERT ChIP-3-F:

GGATAGGCTTCTTTGCTTGC; mTERT ChIP-3-R: TTGATGGTCACAATGCTGGT.

mTERT ChIP-4-F: TTCGTCGTGGACTCTCAGTG; mTERT ChIP-4-R:

GCCACACCTCCCGGTATC. mH19-ICR ChIP-4-F: TCATGGGTCACTCAGGCATA;

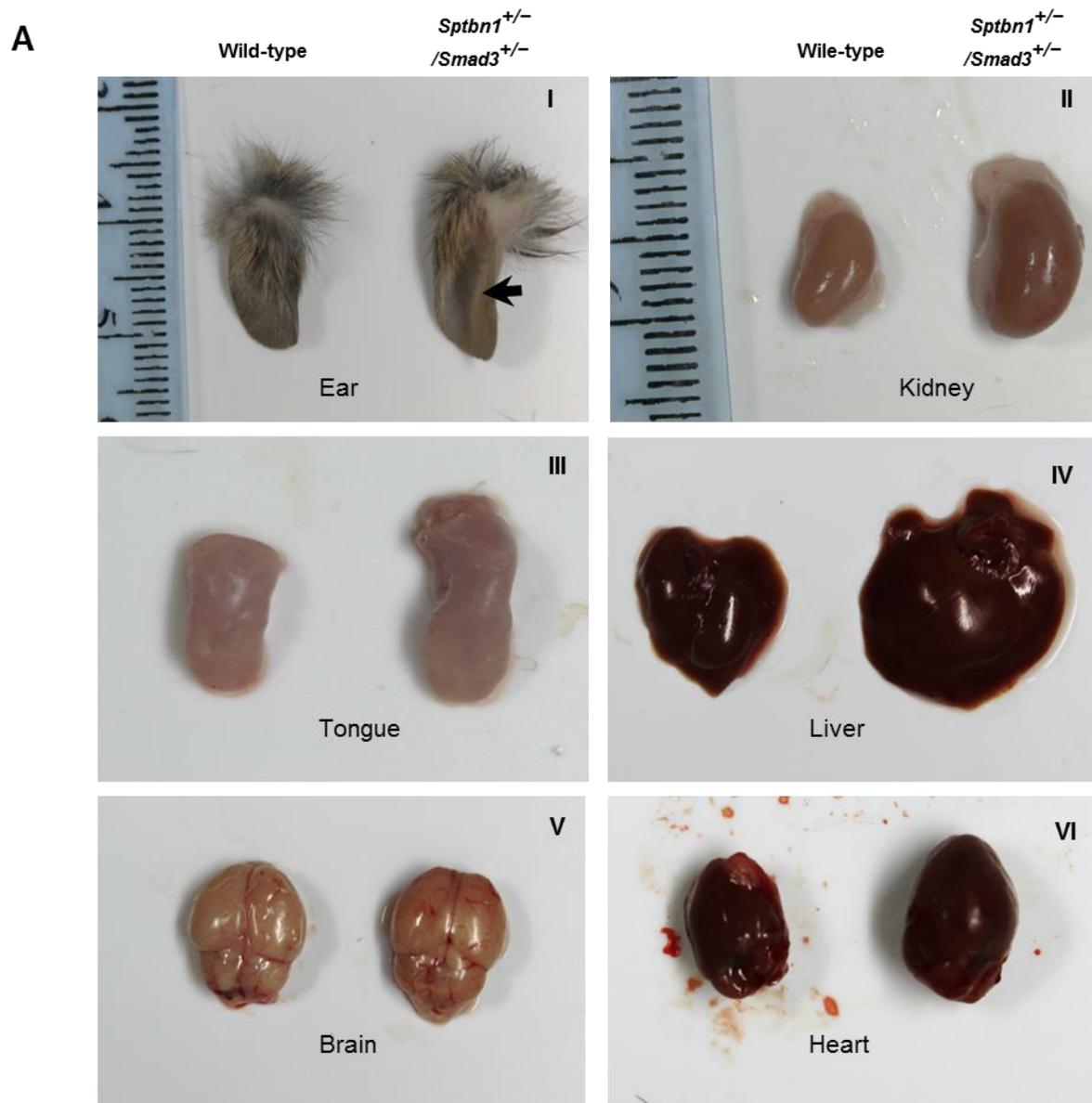
mH19-ICR ChIP-4-R: CGTCTGCCGAGCAATATGTA. hPAI-1 Chip-F:

GTCCTAGGCTTTTTGGGTCA; hPAI-1 Chip-R: ACAATTGAGCAAACCCCAAT. mPAI-1

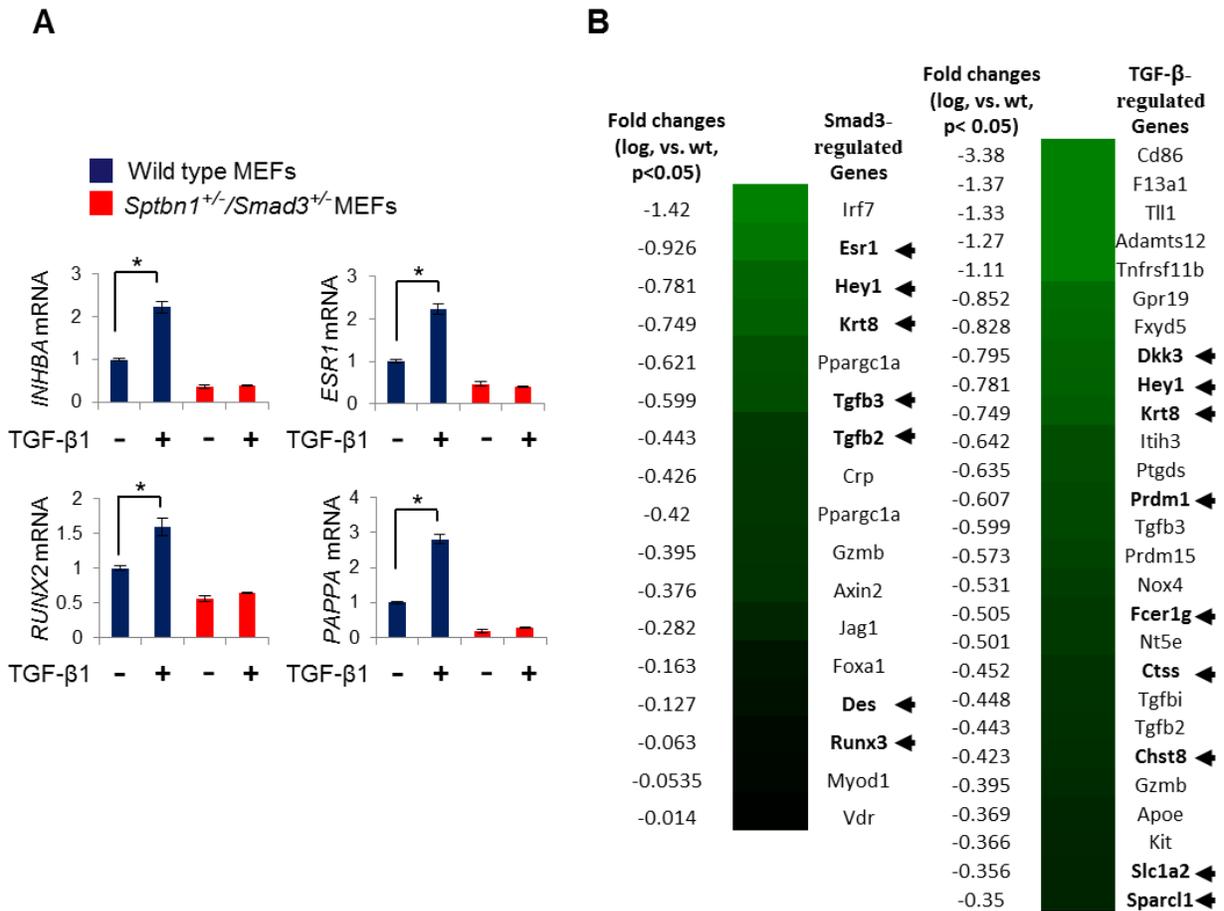
Chip-F: AGGGAACCAGAGTTTGCTCA; mPAI-1 Chip-R: GCCCCACCCACTTTCTAACT.

### **Supplemental References**

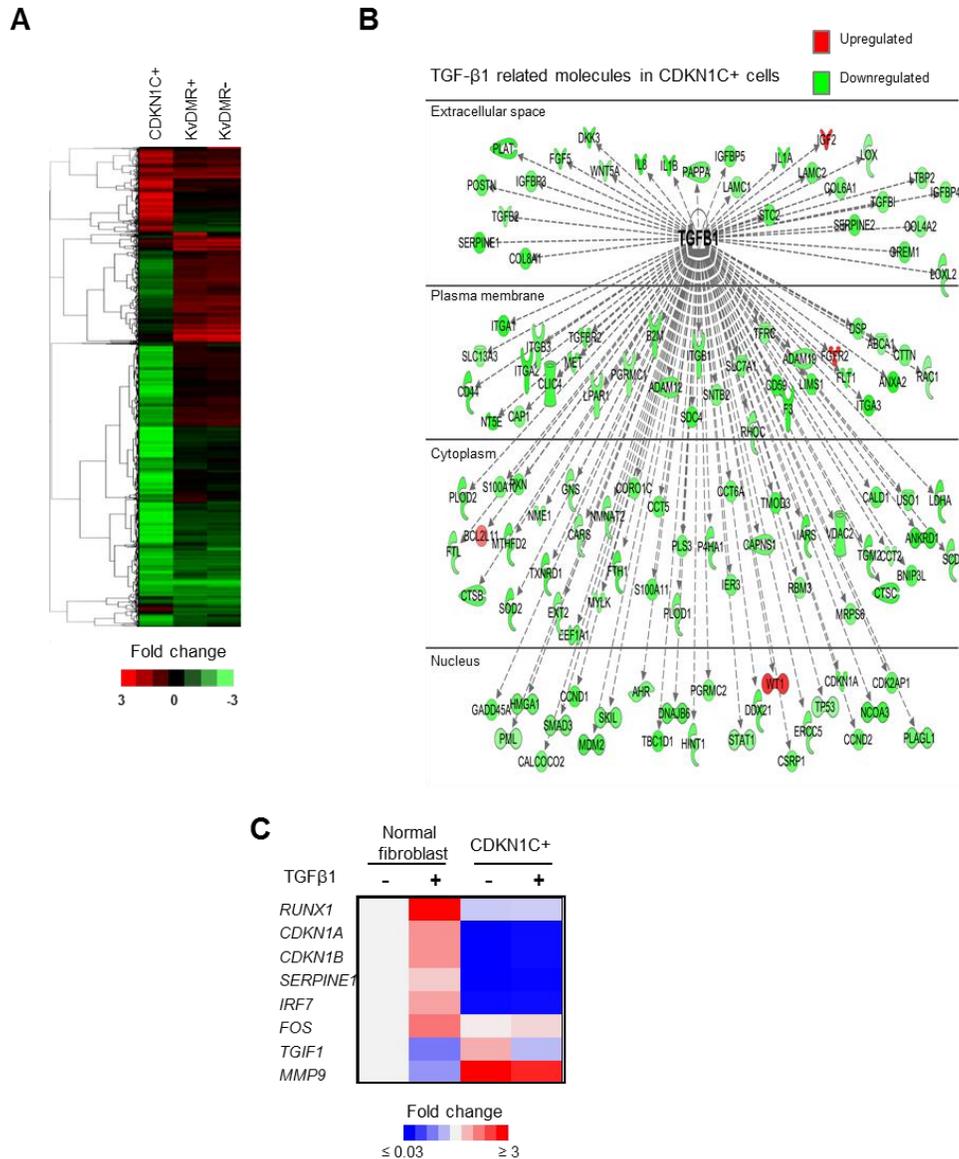
1. Anders S & Huber W (2010) Differential expression analysis for sequence count data. *Genome Biol* 11(10):R106.
2. Marth GT, *et al.* (1999) A general approach to single-nucleotide polymorphism discovery. *Nature genetics* 23(4):452-456.
3. Chen J, *et al.* (2009) Hypoxia-mediated up-regulation of Pim-1 contributes to solid tumor formation. *The American journal of pathology* 175(1):400-411.



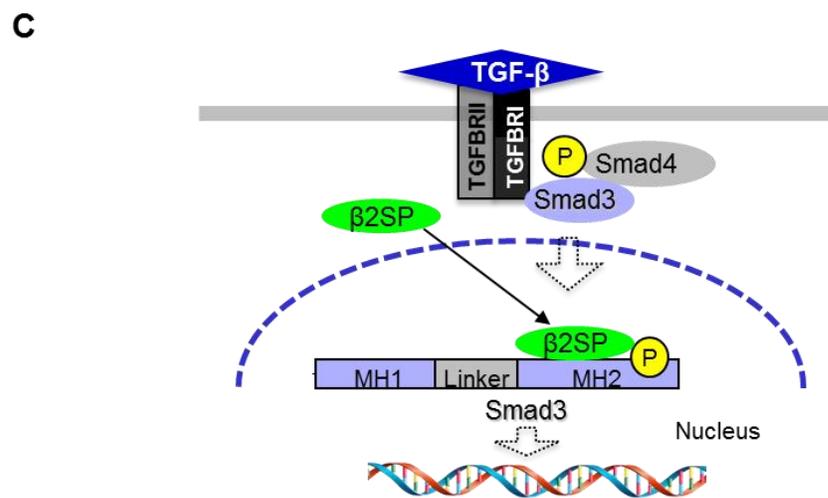
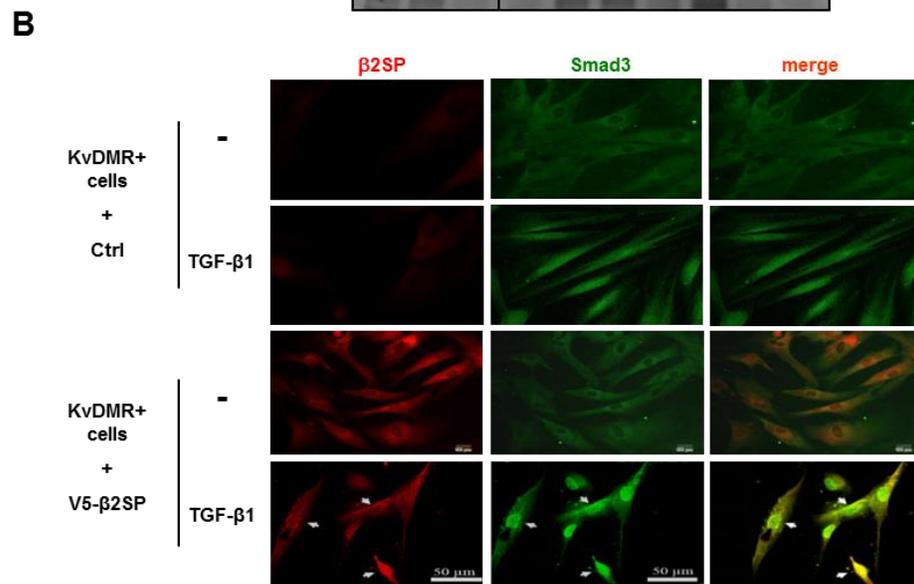
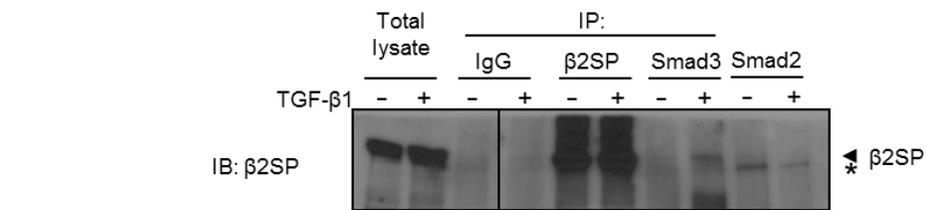
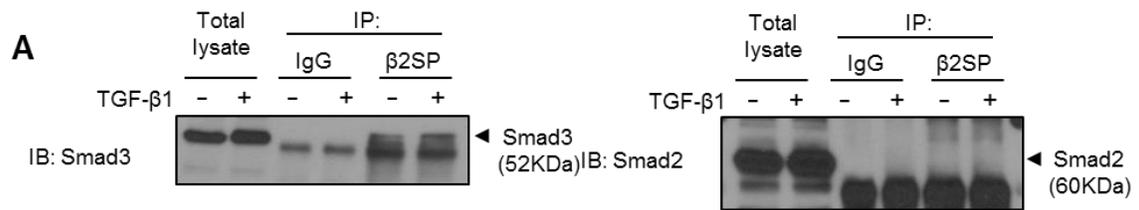
**Supplemental Figure1. *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup> mice phenocopy BWS.** Representative enlargement of organs in *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup> mice. I: ears; II: kidneys; III: Tongues; IV: livers; V: brains; VI: hearts. The arrow indicates an anterior linear ear lobe crease.



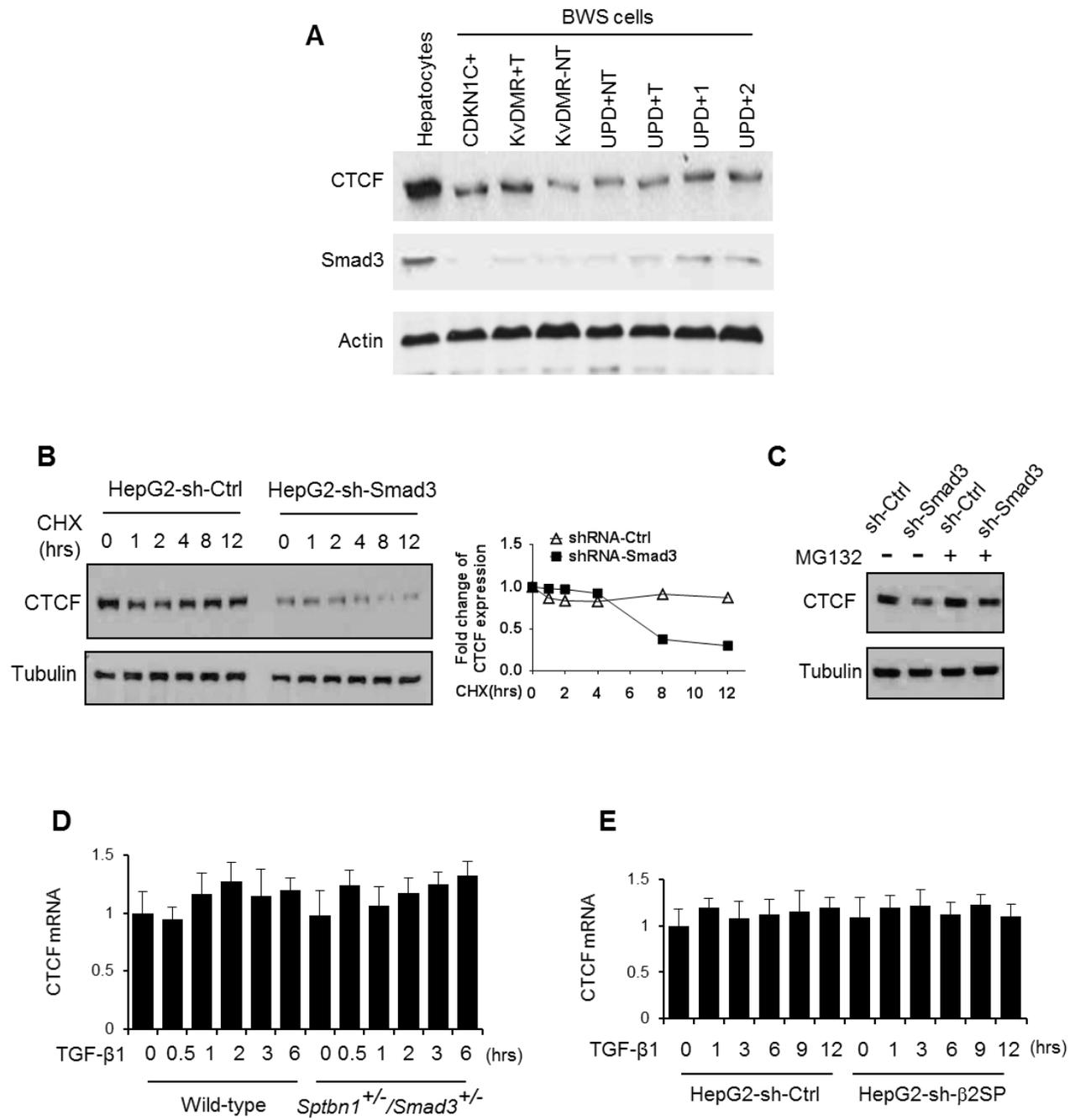
**Supplemental Figure 2. Validation of TGF-β pathway disruption in *Smad3*<sup>-/-</sup> primary keratinocytes.** (A) Q-PCR analysis of defective TGF-β response in *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup> MEFs (n=2). Error bars are shown as standard deviations. Each result shown is representative of three independent experiments. \*:  $P < 0.01$ , Student's t-test. (B) The transcriptome data of two independent *Smad3*<sup>-/-</sup> keratinocytes were downloaded from NCBI's Gene Expression Omnibus (GEO) cohort GSE3228, and further normalized with wild type mouse keratinocytes and then analyzed by IPA. TGF-β/Smad3-regulated genes are detected by fold change cutoff ( $P < 0.05$ ) and presented on heat map.



**Supplemental Figure 3. Analyses of whole-transcriptome sequencing of three BWS cell lines.** (A) The differential gene expressions in three BWS cells were compared to the gene expression in a fibroblast cell line established from a normal person by fold change (significance cutoff,  $P$  value < 0.05). (B) Downregulation of the TGF- $\beta$  pathway in CDKN1C+ BWS cells. The differential gene expression in CDKN1C+ BWS cells was compared to the gene expression in a fibroblast cell line established from normal person and then detected by fold change (cutoff,  $q$  value < 0.3). These data were uploaded to the IPA program. The diagram lists all the TGF- $\beta$ 1-related genes in the input dataset. Red indicates upregulated and green indicates downregulated genes. The arrows point to validated TGF- $\beta$ /Smad3-regulated genes. (C) Q-PCR analysis validated a defective TGF- $\beta$  response in CDKN1C+ cells. Normal fibroblast and BWS CDKN1C+ cells were treated with TGF- $\beta$ 1 for 2 hours. Q-PCR was performed to detect TGF- $\beta$  target gene expression. The result shown is representative of three independent experiments.

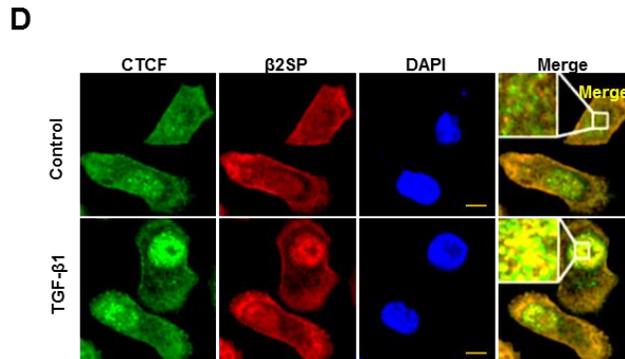
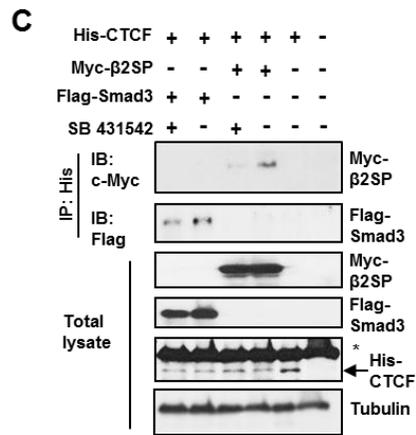
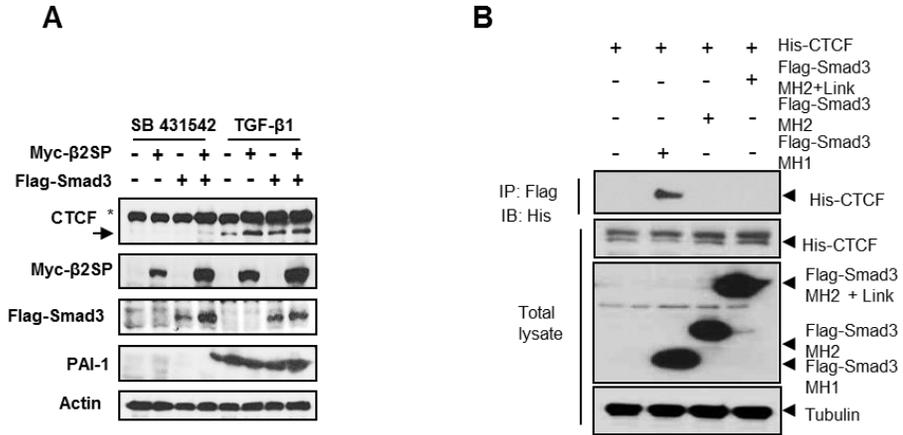


**Supplemental Figure 4.  $\beta$ 2SP interacts with SMAD3.** (A)  $\beta$ 2SP interacts with SMAD3, but not with SMAD2. HepG2 cells were treated with 200 pM TGF- $\beta$  for 2 h. Cell lysates were immunoprecipitated with an anti-Smad3 antibody, an anti-Smad2 antibody or anti- $\beta$ 2SP antibody, and immunoblotted with indicated antibodies. \* designates non-specific bands. (B) Overexpression of  $\beta$ 2SP rescues SMAD3 nuclear translocation in BWS cells. BWS KvDMR+ cells were transfected with full-length V5- $\beta$ 2SP for 24 hours and then were treated with 200 pM TGF- $\beta$ 1 for 2 hours. Immunofluorescent staining was performed to detect  $\beta$ 2SP and Smad3. Scale bars = 50  $\mu$ m. Representative confocal images are shown. Data are representative of two independent experiments (A–B). (C) Proposed model of the role of  $\beta$ 2SP required for Smad3 nuclear translocation in response to TGF- $\beta$ .

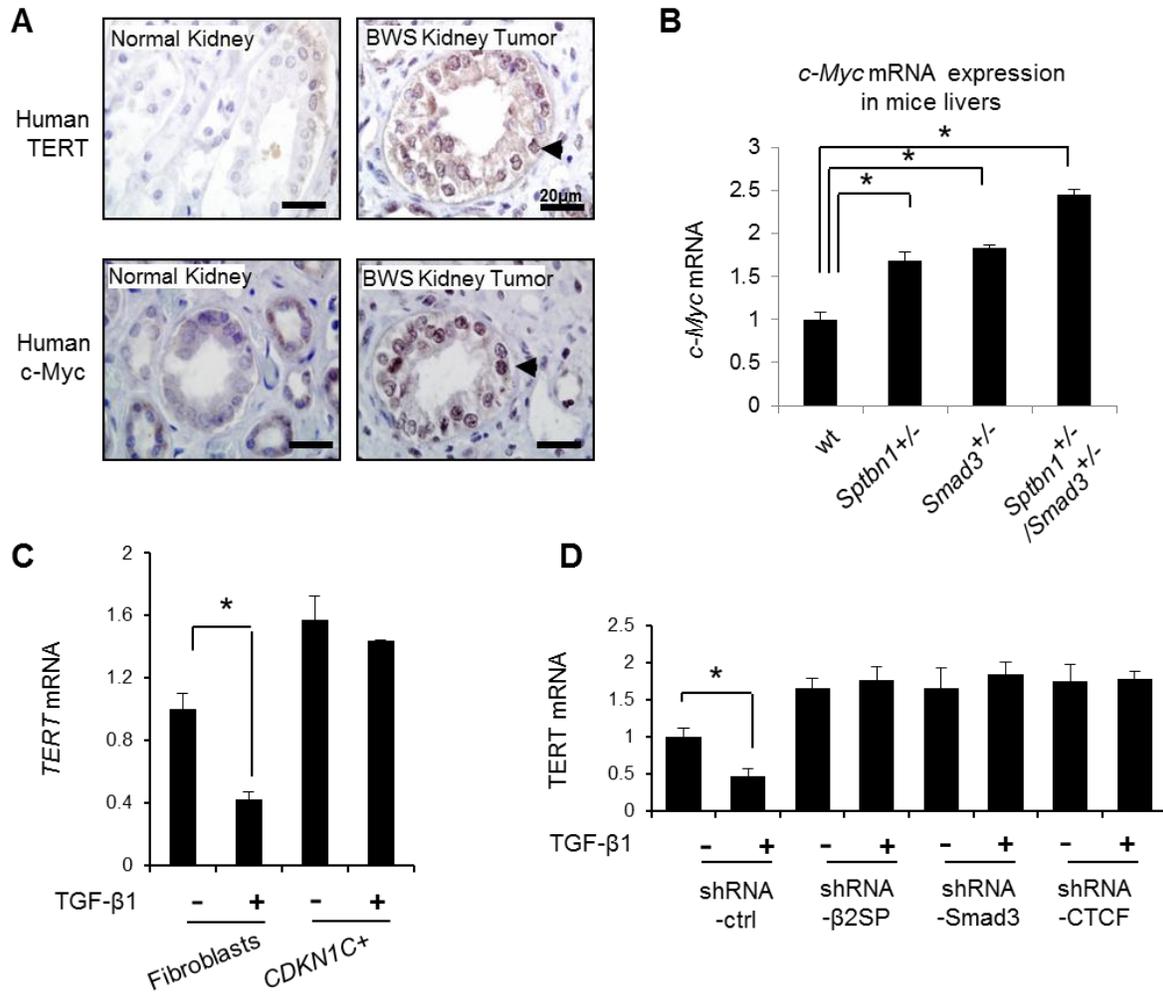


**Supplemental Figure 5. TGF- $\beta$ /Smad3/ $\beta$ 2SP regulates CTCF levels posttranscriptionally.**

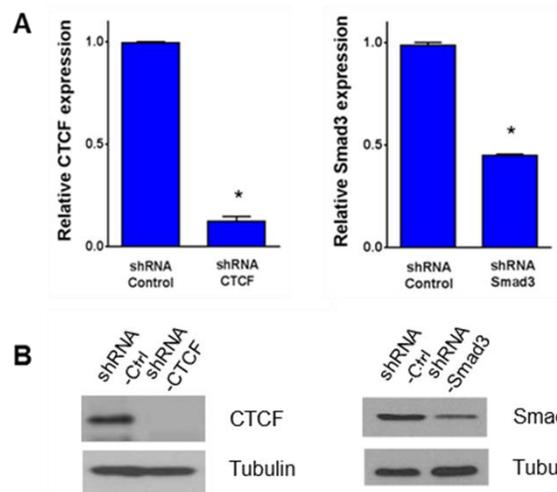
(A) Lower CTCF levels in BWS cells compared with normal hepatocytes. CTCF levels were detected by immunoblotting analysis. Seven human BWS cell lines were developed by Dr. Weksberg (Ontario, Canada). CDKN1C+: CDKN1C mutation, no tumor; KvDMR-: loss of methylation in KCNQ1OT1, no tumor; KvDMR+: loss of methylation in KCNQ1OT1, hepatoblastoma; UPD+NT: uniparental disomy, no tumor; UPD+T: uniparental disomy, hepatoblastoma; UPD+1: uniparental disomy, tongue tissue 1; UPD+2: uniparental disomy, tongue tissue 2. UPD+1 and UPD+2 cell lines were derived from the same case with biopsies at separate sites. KvDMR- and KvDMR+ were derived from normal monozygotic twin (absence of KvDMR molecular defect but it had some clinical signs of BWS) and BWS monozygotic twin with KvDMR molecular defect, respectively. Normal human hepatocytes were received from the Liver Tissue Cell Distribution System, University of Pittsburgh and University of Minnesota. (B) Knock down Smad3 decreased CTCF protein stability. HepG2-sh-Ctrl or HepG2-sh-Smad3 cells were treated with cycloheximide (CHX; 100  $\mu$ g/ml) for the indicated times. The density of CTCF and the integrated optical density were measured. The turnover of CTCF is indicated graphically. (C) Smad3-mediated CTCF downregulation was proteasome-dependent. HepG2-sh-Ctrl or HepG2-sh-Smad3 cells were treated with or without 50  $\mu$ g/ml of MG132 for 6 hours. The cell lysates were then immunoblotted with the indicated antibodies. All blots are representative of experiments performed 3 times (A–C). (D, E) CTCF mRNA level was not affected by TGF- $\beta$ 1 treatment in MEFs (D) or HepG2 cells (E). The *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup> MEFs (D) or  $\beta$ 2SP-knockdown HepG2 cells (E) were treated with 200 pM TGF- $\beta$ 1 for the indicated times. Q-PCR was performed to detect CTCF mRNA expression. (n=3). Each blot is representative of three independent experiments (A, B, D & E), 2(C).



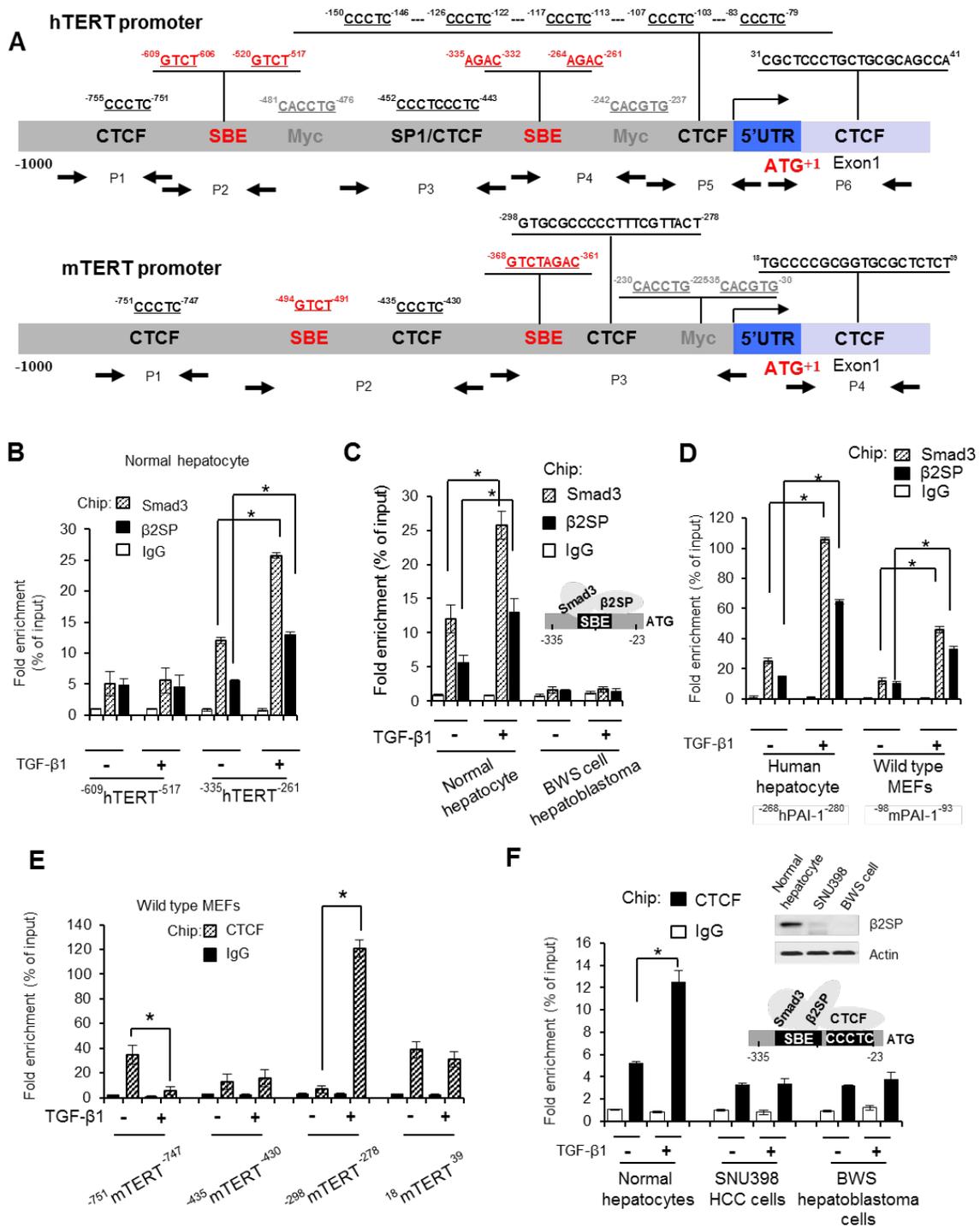
**Supplemental Figure 6.  $\beta$ 2SP/Smad3 interact with CTCF in cell nucleus.** (A)  $\beta$ 2SP and/or Smad3 increased CTCF levels in a TGF- $\beta$ -dependent manner. HepG2 cells were co-transfected with the indicated plasmids and were treated with TBR1 inhibitor SB431542 (5  $\mu$ M) overnight or with TGF- $\beta$ 1 (200 pM) for 2 h. (B) CTCF interacts with Smad3 MH1 domain. SNU398 cells were co-transfected with indicated plasmids. Cell lysates were immunoprecipitated with an anti-His antibody and immunoblotted with indicated antibodies. (C) Interaction of  $\beta$ 2SP/Smad3 with CTCF is TGF- $\beta$  dependent. SNU398 cells were co-transfected with the indicated plasmids and were treated with 5  $\mu$ M SB431542 overnight. Cell lysates were immunoprecipitated with an anti-histidine antibody and immunoblotted with indicated antibodies. Each blot is representative of two independent experiments (A–C). (D) Nuclear translocation of  $\beta$ 2SP and CTCF upon treatment of TGF- $\beta$ . SNU475 cells were treated with 200 pM TGF- $\beta$ 1 for 2 h. Immunofluorescent staining was performed to detect  $\beta$ 2SP and CTCF. Scale bars, 20  $\mu$ m.



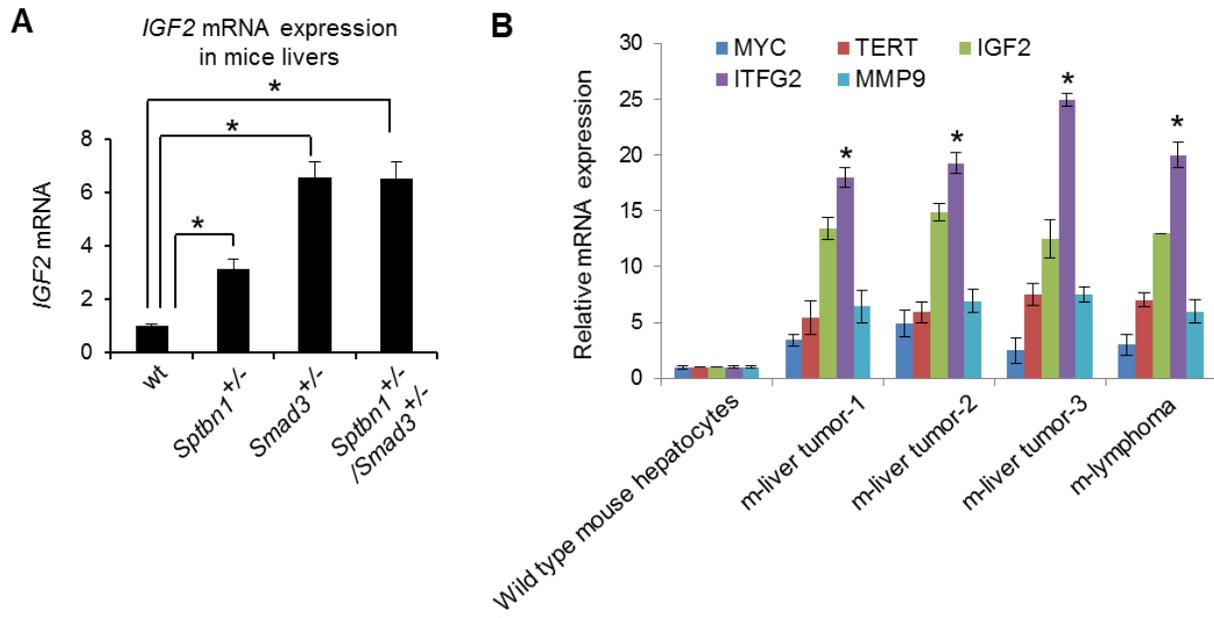
**Supplemental Figure 7. The expression of *TERT*, *c-Myc* and *IGF2* are increased in human BWS-associated tumors, *TGF- $\beta$*  defective mouse livers and cell lines established from individual tumors in *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup> mice. (A) Increased levels of TERT and c-Myc in kidney tumors of BWS patients were observed. Representative immunohistochemical staining of human TERT and c-Myc in normal kidney or in kidney tumors of BWS patients. (B) Increased mRNA expressions of *c-Myc* in *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup>, *Smad3*<sup>+/-</sup>, and *Sptbn1*<sup>+/-</sup> mouse livers. (C) Reduction of TGF- $\beta$ -induced *TERT* expression levels were observed in human normal fibroblasts but not CDKN1C<sup>+</sup> BWS cells. The cells were treated with 200 pM TGF- $\beta$  for 2 h. (D) The *TERT* mRNA level was increased in  $\beta$ 2SP, Smad3 or CTCF knockdown cells compared with control cells. The cells were treated with 200 pM TGF- $\beta$  for 2 h. Error bars are shown as standard deviations. Each result shown is representative of three independent experiments (B–D). \*:  $P < 0.001$ , one-way ANOVA with post-hoc Bonferroni test**



**Supplemental Figure 8. Knock down CTCF or Smad3 in SNU398 cells.** SNU398 cells were infected with lentivirus-mediated CTCF-shRNA, Smad3-shRNA or control-shRNA. **(A)** Q-PCR and **(B)** Western blot analyses were performed to detect the expression of CTCF and Smad3 in SNU398 cells. Results are the average of three independent experiments and are presented as mean  $\pm$  SD in (A). \*:  $P < 0.01$ , Student's t-test.



**Supplemental Figure 9.  $\beta$ 2SP, Smad3, and CTCF bind on *TERT* Promoter Region.** (A) Diagrammatic representation shows the potential SBE motifs, CTCF binding motifs, and Myc binding motifs on the human or mouse *TERT* promoter-exon1 region. (B) TGF- $\beta$  increases  $\beta$ 2SP/Smad3 binding activities on the sites ( $^{-335}\text{hTERT}^{-261}$  but not  $^{-609}\text{hTERT}^{-517}$ ) on human *TERT* promoter region. Genomic DNA was isolated from normal hepatocytes treated with TGF- $\beta$ 1 (200 pM) for 2 h. (C) TGF- $\beta$ 1 treatment increases  $\beta$ 2SP and SMAD3 binding on the *TERT* promoter in normal hepatocytes but not in BWS KvDMR+T hepatoblastoma cells. The cells were treated with TGF- $\beta$ 1 for 2 h. (D) TGF- $\beta$ 1 increases  $\beta$ 2SP/Smad3 binding activities on human and mouse *PAI-1* promoter. Genomic DNA was isolated from normal hepatocytes and wild type MEFs. ChIP assays were performed and enrichment of  $\beta$ 2SP or Smad3 transcripts was measured by quantitative PCR for (B) to (D). (E) TGF- $\beta$  increases CTCF binding activities on the sites ( $^{-298}\text{GTGCGCCCCCTTTCGTTAT}^{-278}$ , but not on  $^{-751}\text{CCCTC}^{-747}$  or on  $^{-435}\text{CCCTC}^{-430}$ ) on mouse *TERT* promoter region. ChIP assays were performed to detect the binding activity of CTCF on the mouse *TERT* promoter region. Genomic DNA was isolated from wild type MEFs treated with TGF- $\beta$ 1. (F) Binding ability of CTCF on *hTERT* promoter region is  $\beta$ 2SP dependent. Normal human hepatocytes or SNU398 or BWS KvDMR+T cells were treated with TGF- $\beta$ 1 for 2 h and ChIP assays were performed. Enrichment of CTCF transcripts was measured by quantitative PCR. Error bars are shown as standard deviations. Each result shown is representative of three independent experiments (B-F) \*:  $P < 0.001$ , one-way ANOVA with post-hoc Bonferroni test.



**Supplemental Figure 10. Increased levels of IGF2 in livers and tumors in**

*Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup> mice. (A) Increased mRNA expressions of IGF2 in *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup>,

*Smad3*<sup>+/-</sup>, and *Sptbn1*<sup>+/-</sup> mouse livers. (B) Increased mRNA expressions of c-Myc, TERT,

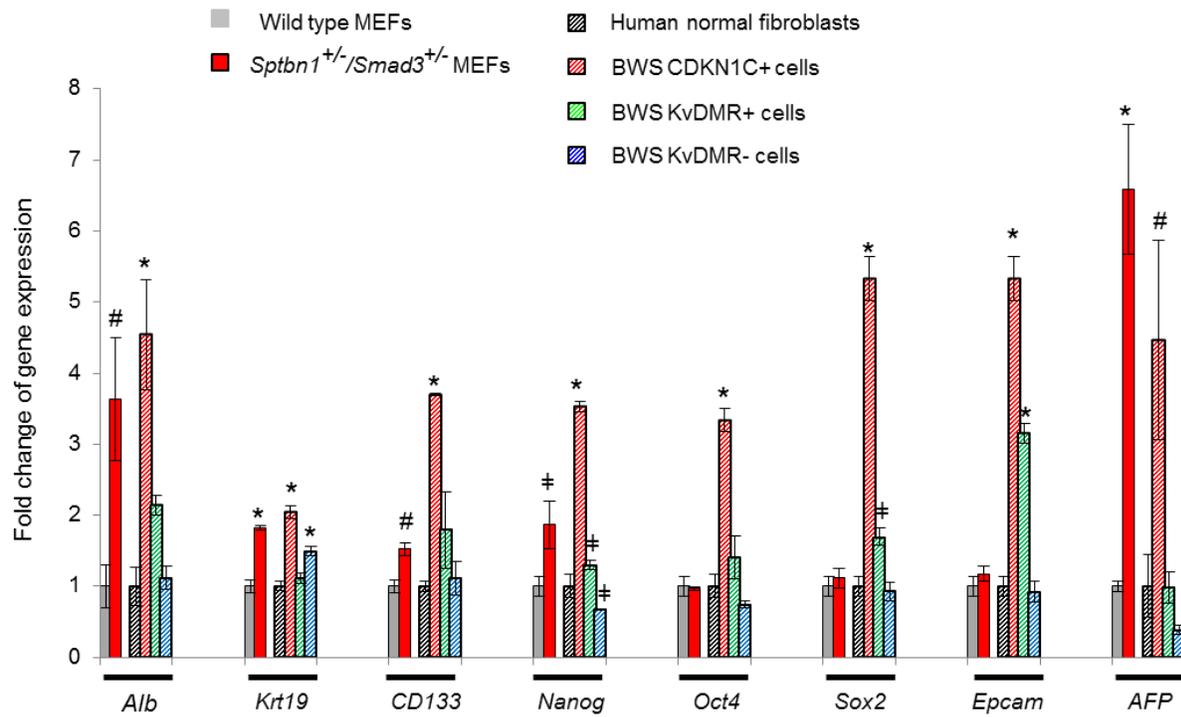
IGF2, ITFG2 and MMP9 in mouse tumor cell lines. 4 cell lines were established from individual

tumors (three hepatocarcinomas and one lymphoma) in *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup> mice. Results are

the average of three independent experiments and are presented as mean ± SD (A and B). \*: *P* <

0.001, versus wild type mouse hepatocytes or wild type MEFs, one-way ANOVA with post-hoc

Bonferroni test.



**Supplemental Figure 11.** Increased mRNA expression levels of stemness genes in MEFs and BWS cell lines. Q-PCR was performed to detect gene expression. Error bars are shown as standard deviations. Results are the average of three independent experiments and are presented as mean  $\pm$  SD. †:  $P < 0.05$ , #:  $P < 0.01$ , \*:  $P < 0.001$ , versus wild type MEFs (Student's t-test) or human normal fibroblasts (one-way ANOVA with post-hoc Bonferroni test).

**Supplemental Table 1. Classification of tumors in *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup>, *Sptbn1*<sup>+/-</sup> and *Smad3*<sup>+/-</sup> mice.**

Site	Tumor Categories	<i>Sptbn1</i> <sup>+/-</sup> / <i>Smad3</i> <sup>+/-</sup>		<i>Sptbn1</i> <sup>+/-</sup>		<i>Smad3</i> <sup>+/-</sup>	
		No. of mice	Incidence (%)	No. of mice	Incidence (%)	No. of mice	Incidence (%)
Small Intestine	Adenocarcinoma	5	55.56	2	4.65	0	0
Liver	Hepatocellular Carcinoma	5	55.56	16	37.21	1	3.33
Pancreas	Adenocarcinoma	1	11.11	0	0.00	0	0
Head & Neck	Squamous Cell Carcinoma	1	11.11	0	0.00	0	0
Lung	Adenocarcinoma	5	55.56	2	4.65	1	3.33
Colon	Adenocarcinoma	2	22.22	1	2.33	0	0
Thymus	Thymoma	2	22.22	0	0.00	0	0
Mediastinal Mass	Sarcoma	2	22.22	1	2.33	0	0
Spleen	Lymphoma	1	11.11	2	4.65	1	3.33
Kidney	Clear Cell Carcinoma	2	22.22	3	6.98	0	0
Adrenal	Adrenocortical Carcinoma	2	22.22	0	0.00	0	0
Skin	Squamous Cell Carcinoma	1	11.11	0	0.00	0	0
Breast	Adenoma	2	22.22	1	2.33	0	0
Lacrimal Gland	Adenocarcinoma	1	11.11	0	0.00	0	0
Optic Nerve	Glioma	1	11.11	0	0.00	0	0

Total *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup> mice, n=15; total mice *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup> with tumors, n=11; mice with multiple tumors, n=9.

Total *Sptbn1*<sup>+/-</sup> mice n=43; total mice *Sptbn1*<sup>+/-</sup> with tumors, n=19; mice with multiple tumors, n=4.

Total *Smad3*<sup>+/-</sup> mice n=30; total mice *Smad3*<sup>+/-</sup> with tumors, n=3; mice with multiple tumors n=0.

**Supplemental Table 2. Sporadic tumor development in *Sptbn1*<sup>+/-</sup>/*Smad3*<sup>+/-</sup>, *Sptbn1*<sup>+/-</sup> and *Smad3*<sup>+/-</sup> mice.**

Genotype	Mouse number	Gender	Thymoma	Head & Neck cancer	Glioma	Lacrimal gland cancer	Skin cancer	Adrenocortical carcinoma	Pancreatic cancer	Kidney cancer	Sarcoma	Breast cancer	Colon adenocarcinoma	Small Intestine cancer	Lymphoma	Lung cancer	Hepatocellular carcinoma
<i>Sptbn1</i> <sup>+/-</sup> / <i>Smad3</i> <sup>+/-</sup>	1	F		+	+				+					+		+	+
	2	F											+			+	+
	3	M	+			+								+			+
	4	M									+			+			+
	5	M		+			+							+			+
	6	F	+									+		+			+
	7	M						+		+							+
	8	F											+				+
	9	F						+		+				+			
	10	M														+	
	11	M										+					
		12-15*	F:2 M:2														
<i>Sptbn1</i> <sup>+/-</sup>	1	M												+	+	+	+
	2	F												+			+
	3	M								+	+		+				+
	4	M															+
	5	M															+
	7	F															+
	7	F															+
	8	F															+
	9	F															+
	10	M															+
	11	F															+
	12	M															+
	13	F															+
	14	M															+
	15	M															+
	16	M															+
	17	F								+		+					+
	18	M								+							
	19	F														+	
	20-43*	F:10 M: 14															
<i>Smad3</i> <sup>+/-</sup>	1	M															+
	2	F														+	
	3	F													+		
	4-30*	F:14 M:13															

\*: Mice without tumor.

**Supplemental Table 3. Somatic mutations of Smad3 and  $\beta$ 2SP identified on COSMIC and LIHC TCGA.**

COSMIC						
Gene Name	Sample ID	AA Mutation	CDS Mutation	Primary Tissue	Histology	Somatic Status
SMAD3	1688022	p.A349S	c.1045G>T	Autonomic ganglia	Neuroblastoma	Confirmed Somatic
SMAD3	1735399	p.V331I	c.991G>A	Central nervous system	Primitive neuroectodermal tumour-medulloblastoma	Confirmed Somatic
SMAD3	1783454	p.R268C	c.802C>T	Endometrium	Carcinoma	Confirmed Somatic
SMAD3	1783371	p.R295W	c.883C>T	Endometrium	Carcinoma	Confirmed Somatic
SMAD3	1783520	p.V331I	c.991G>A	Endometrium	Carcinoma	Confirmed Somatic
SMAD3	1783419	p.S425Y	c.1274C>A	Endometrium	Carcinoma	Confirmed Somatic
SMAD3	1897529	p.Y226fs*1	c.676_677insA	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SMAD3	1766754	p.R90H	c.269G>A	Large intestine	Carcinoma	Confirmed Somatic
SMAD3	1651648	p.E178D	c.534G>T	Large intestine	Carcinoma	Confirmed Somatic
SMAD3	1651567	p.Y237C	c.710A>G	Large intestine	Carcinoma	Confirmed Somatic
SMAD3	1766767	p.N276I	c.827A>T	Large intestine	Carcinoma	Confirmed Somatic
SMAD3	1766809	p.R287W	c.859C>T	Large intestine	Carcinoma	Confirmed Somatic
SMAD3	1766754	p.R368Q	c.1103G>A	Large intestine	Carcinoma	Confirmed Somatic
SMAD3	1766776	p.R373H	c.1118G>A	Large intestine	Carcinoma	Confirmed Somatic
SMAD3	1651670	p.G379A	c.1136G>C	Large intestine	Carcinoma	Confirmed Somatic
SMAD3	1235044	p.P393L	c.1178C>T	Large intestine	Carcinoma	Confirmed Somatic
SMAD3	1651671	p.P393L	c.1178C>T	Large intestine	Carcinoma	Previously Reported
SMAD3	2067140	p.R142H	c.425G>A	Liver	Carcinoma	Confirmed Somatic
SMAD3	1863736	p.E239V	c.716A>T	Lung	Carcinoma	Confirmed Somatic
SMAD3	1780235	p.R243C	c.727C>T	Lung	Carcinoma	Confirmed Somatic
SMAD3	1780064	p.P253S	c.757C>T	Lung	Carcinoma	Confirmed Somatic
SMAD3	1780171	p.R268H	c.803G>A	Lung	Carcinoma	Confirmed Somatic
SMAD3	1860019	p.E149G	c.446A>G	Oesophagus	Carcinoma	Confirmed Somatic
SMAD3	1434856	p.S423G	c.1267A>G	Ovary	Carcinoma	Confirmed Somatic
SMAD3	1473141	p.S422F	c.1265C>T	Pancreas	NS	Confirmed Somatic
SMAD3	1797209	p.A329S	c.985G>T	Prostate	Carcinoma	Confirmed Somatic
SMAD3	2013572	p.R373C	c.1117C>T	Skin	Malignant melanoma	Confirmed Somatic

SMAD3	1542864	p.S425C	c.1274C>G	Upper aerodigestive tract	Carcinoma	Confirmed Somatic
SPTBN1	1659943	p.A682V	c.2045C>T	Breast	Carcinoma	Confirmed Somatic
SPTBN1	1659947	p.H1760P	c.5279A>C	Breast	Carcinoma	Confirmed Somatic
SPTBN1	1779256	p.E1646K	c.4936G>A	Breast	Carcinoma	Confirmed Somatic
SPTBN1	1779322	p.R42W	c.124C>T	Breast	Carcinoma	Confirmed Somatic
SPTBN1	1430345	p.R107*	c.319C>T	Breast	Carcinoma	Confirmed Somatic
SPTBN1	1783319	p.A2326T	c.6976G>A	Endometrium	Carcinoma	Confirmed Somatic
SPTBN1	1783328	p.R1258G	c.3772A>G	Endometrium	Carcinoma	Confirmed Somatic
SPTBN1	1783387	p.D1813N	c.5437G>A	Endometrium	Carcinoma	Confirmed Somatic
SPTBN1	1783412	p.E475K	c.1423G>A	Endometrium	Carcinoma	Confirmed Somatic
SPTBN1	1783443	p.A395T	c.1183G>A	Endometrium	Carcinoma	Confirmed Somatic
SPTBN1	1783467	p.V1946I	c.5836G>A	Endometrium	Carcinoma	Confirmed Somatic
SPTBN1	1783524	p.F2236I	c.6706T>A	Endometrium	Carcinoma	Confirmed Somatic
SPTBN1	1945136	p.G246S	c.736G>A	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SPTBN1	1945111	p.L1973I	c.5917C>A	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SPTBN1	1945150	p.T2320M	c.6959C>T	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SPTBN1	1945173	p.K497R	c.1490A>G	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SPTBN1	1945181	p.K401T	c.1202A>C	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SPTBN1	1945190	p.N1226S	c.3677A>G	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SPTBN1	1998437	p.V1860M	c.5578G>A	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SPTBN1	1630874	p.R681W	c.2041C>T	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SPTBN1	1998456	p.F694Y	c.2081T>A	Haematopoietic and lymphoid	Lymphoid neoplasm	Confirmed Somatic
SPTBN1	1921467	p.R681W	c.2041C>T	Haematopoietic and lymphoid	Lymphoid neoplasm	Previously Reported
SPTBN1	1766747	p.D161Y	c.481G>T	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1766747	p.L2084S	c.6251T>C	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1766748	p.H1634R	c.4901A>G	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1766750	p.R384W	c.1150C>T	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1766751	p.A2240T	c.6718G>A	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1766769	p.R913H	c.2738G>A	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1766789	p.R1375Q	c.4124G>A	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1766793	p.W970R	c.2908T>C	Large intestine	Carcinoma	Confirmed Somatic

SPTBN1	1766805	p.Q1440H	c.4320A>C	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1998441	p.R1739Q	c.5216G>A	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1998483	p.N1527I	c.4580A>T	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1998442	p.L1948F	c.5842C>T	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1651542	p.E622A	c.1865A>C	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1651648	p.E1533*	c.4597G>T	Large intestine	Carcinoma	Confirmed Somatic
SPTBN1	1651673	p.R107*	c.319C>T	Large intestine	Carcinoma	Previously Reported
SPTBN1	2067095	p.S2245A	c.6733T>G	Liver	Carcinoma	Confirmed Somatic
SPTBN1	2067170	p.L1355V	c.4063C>G	Liver	Carcinoma	Confirmed Somatic
SPTBN1	1765225	p.K593fs*1 1	c.1777_1781delAAGT T	Lung	Carcinoma	Confirmed Somatic
SPTBN1	1434855	p.I1636M	c.4908C>G	Ovary	Carcinoma	Confirmed Somatic
SPTBN1	1474856	p.L962F	c.2884C>T	Ovary	Carcinoma	Confirmed Somatic
SPTBN1	1475027	p.L1314*	c.3941T>A	Ovary	Carcinoma	Confirmed Somatic
SPTBN1	2068114	p.R144W	c.430C>T	Pancreas	Carcinoma	Confirmed Somatic
SPTBN1	1998469	p.M133T	c.398T>C	Prostate	Carcinoma	Confirmed Somatic
SPTBN1	1604874	p.M526I	c.1578G>A	Skin	Carcinoma	Confirmed Somatic
SPTBN1	1604872	p.R972Q	c.2915G>A	Skin	Carcinoma	Confirmed Somatic
SPTBN1	2013568	p.R1915C	c.5743C>T	Skin	Malignant melanoma	Confirmed Somatic
SPTBN1	1673883	p.R1915C	c.5743C>T	Skin	Malignant melanoma	Previously Reported
<b>LIHC</b>						
<b>TCGA</b>						
<b>Gene Name</b>	<b>Tumor_Sample_Barcode</b>	<b>AA Change</b>	<b>Chrom Change</b>	<b>Primary Tissue</b>	<b>Variant</b>	<b>Mutation Status</b>
SMAD3	TCGA-DD-A11C-01A-11D-A12Z-10	p.F376L	c.T1126C	Liver	Missense Mutation	Somatic
SMAD3	TCGA-DD-A1EL-01A-11D-A152-10	p.L42F	c.C124T	Liver	Missense Mutation	Somatic
SMAD3	TCGA-DD-A39Y-01A-11D-A20W-10	p.T392fs	c.1175delC	Liver	Frame_Shift_Del	Somatic
SMAD3	TCGA-G3-A6UC-01A-21D-A33K-10	p.G25V	c.G74T	Liver	Missense Mutation	Somatic
SMAD3	TCGA-RC-A6M4-01A-11D-A32G-10	p.K333E	c.A997G	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-BW-A5NP-01A-11D-A27I-10	p.N32Y	c.A94T	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-DD-A119-01A-11D-A12Z-10	p.E33D	c.G99T	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-DD-A3A1-01A-11D-A20W-10	p.L388R	c.T1163G	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-FV-A3R2-01A-11D-A22F-10	p.E489K	c.G1465A	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-G3-A25Y-01A-11D-A16V-10	p.G562C	c.G1684T	Liver	Missense Mutation	Somatic

SPTBN1	TCGA-FV-A4ZQ-01A-11D-A25V-10	p.N1106S	c.A3317G	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-DD-A1EH-01A-11D-A12Z-10	p.R1141W	c.C3421T	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-QA-A7B7-01A-11D-A32G-10	p.A1220V	c.C3659T	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-BC-A112-01A-11D-A12Z-10	NA	c.4495-2A>G	Liver	Splice_Site	Somatic
SPTBN1	TCGA-UB-A7MB-01A-11D-A33Q-10	p.A1645E	c.C4934A	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-DD-A4NG-01A-11D-A27I-10	p.L1783R	c.T5348G	Liver	Missense Mutation	Somatic
SPTBN1	TCGA-EP-A265-01A-11D-A16V-10	p.W653R	c.T1957C	Liver	Missense Mutation	Somatic
<b>MD Anderson</b>						
<b>Gene Name</b>	<b>Tumor_Sample_Barcode</b>	<b>AA Change</b>	<b>Chrom Change</b>	<b>Primary Tissue</b>	<b>Variant Classification</b>	<b>Mutation Status</b>
SPTBN1	9194MDACC	p.D1089Y	c.G3265T	Liver	Missense_Mutation	Somatic

**Supplemental Table 4. Mutual exclusivity of co-occurrence of SPTBN1 with SMAD2, SMAD3 or SMAD4 in TCGA dataset.\***

Co-occur	Cancer type	Cohort	Gene A	Gene B	p-Value <sup>†</sup>	Log Odds Ratio	Association
SPTBN1+SMAD2	Colorectal Adenocarcinoma	Genentech	SPTBN1	SMAD2	0.016 <sup>‡</sup>	2.501	significant tendency towards co-occurrence
		TCGA	SPTBN1	SMAD2	0.093	1.65	association towards co-occurrence
	Stomach Adenocarcinoma	TCGA	SPTBN1	SMAD2	0.444	0.659	association towards co-occurrence
	Uterine Corpus Endometrioid Carcinoma	TCGA	SPTBN1	SMAD2	0.495	0.508	association towards co-occurrence
SPTBN1+SMAD3	Cervical Squamous Cell Carcinoma	TCGA	SPTBN1	SMAD3	0.001 <sup>‡</sup>	>3	significant tendency towards co-occurrence
	Kidney Chromophobe	TCGA	SPTBN1	SMAD3	0.015 <sup>‡</sup>	>3	significant tendency towards co-occurrence
	Colorectal Adenocarcinoma	TCGA	SPTBN1	SMAD3	0.256	1.451	association towards co-occurrence
	Head and Neck Squamous Cell Carcinoma	TCGA	SPTBN1	SMAD3	0.077	>3	association towards co-occurrence
	Nasopharyngeal Carcinoma	Singapore	SPTBN1	SMAD3	0.036 <sup>‡</sup>	>3	significant tendency towards co-occurrence
	Uterine Corpus Endometrioid Carcinoma	TCGA	SPTBN1	SMAD3	0.102	1.601	association towards co-occurrence
SPTBN1+SMAD4	Cervical Squamous Cell Carcinoma	TCGA	SPTBN1	SMAD4	0.964	<-3	association towards co-occurrence
	Colorectal Adenocarcinoma	Genentech	SPTBN1	SMAD4	0.281	0.762	association towards co-occurrence
	Stomach Adenocarcinoma	Pfizer/UHK	SPTBN1	SMAD4	0.144	2.543	association towards co-occurrence
	Uterine Corpus Endometrioid Carcinoma	TCGA	SPTBN1	SMAD4	0.053	2.12	association towards co-occurrence

\*: Analyzed by cBio portal complements existing tools (<http://cbioportal.org>).

<sup>†</sup>: Derived from Fisher Exact Text.

<sup>‡</sup>: p-Value < 0.05: significant tendency towards co-occurrence.

Log Odds Ratio > 0: association towards co-occurrence; Log Odds Ratio ≤ 0: association towards mutual exclusivity.

**Supplemental Table 5. Top 5 predicted activation and inactivation diseases/functions in MEFs.**

	<i>Sptbn1</i> <sup>+/-</sup> / <i>Smad3</i> <sup>+/-</sup>	z-score	<i>Sptbn</i> <sup>+/-</sup>	z-score	<i>Sptbn1</i> <sup>-/-</sup>	z-score	<i>Smad3</i> <sup>+/-</sup>	z-score
Activation:	Cell movement	3.758	Organism death	2.651	Hypertrophy	2.425	Organism death	3.579
	Leukocyte migration	3.46	Hypertrophy of heart	2.405	Size of infarct	2.401	Synthesis of polysaccharide	2.983
	Migration of cells	3.422	Cardiac fibrosis	2.126	Size of lesion	2.209	Metabolism of carbohydrate	2.718
	Cell movement of leukocytes	3.197	Abnormal morphology of body cavity	1.982	Lesion formation	2.209	Blood pressure	2.351
	Chemotaxis of cells	3.073	Quantity of vitamin	1.98	Heart Disease	2.025	Multiple congenital anomalies	1.982
Inactivation:	Organism death	-3.816	Activation of blood cells	-3.89	Damage of connective tissue	-2.433	Size of body	-3.388
	Infection of mammalia	-2.016	Activation of leukocytes	-3.674	Accumulation of phagocytes	-2.348	Metabolism of terpenoid	-2.781
	Parasitic infection	-1.969	Activation of cells	-3.659	Damage of cartilage tissue	-2.219	Analgesia	-2.578
	Inflammation of organ	-1.421	Activation of myeloid cells	-3.238	Accumulation of myeloid cells	-1.924	Antinociception	-2.57
	Nephritis	-1.406	Adhesion of blood cells	-3.213	Cell movement of macrophages	-1.626	Infiltration of cells	-2.437

**Supplemental Table 6. Top 5 predicted activation and inactivation diseases/functions in BWS cell lines.**

	CDKN1C+	z-score	KvDMR+	z-score	KvDMR-	z-score
Activation:	Organism death	9.611	Fertility	3.209	<b>Size of body</b>	2.428
	Growth failure	5.509	Proliferation of embryonic cells	2.263	Proliferation of embryonic cells	2.244
	Anemia	3.911	Synthesis of DNA	2.233	Lymphangiogenesis	1.969
	Heart disease	3.709	Cell surface receptor linked signal transduction	2.156	Cell surface receptor linked signal transduction	1.941
	<b>Tumorigenesis of malignant tumor</b>	3.273	<b>Size of body</b>	1.93	Synthesis of DNA	1.898
Inactivation:	Size of body	-6.418	Organism death	-3.226	Organism death	-4.163
	Cell movement	-5.757	Morphology of digestive system	-2.771	Perinatal death	-2.955
	Viral infection	-5.111	Adhesion of blood cells	-2.478	Damage of lung	-2.954
	Cell viability	-5.062	Perinatal death	-2.464	Adhesion of blood cells	-2.548
	Cell survival	-4.913	Congenital anomaly of musculoskeletal system	-2.423	Hypoplasia	-2.507