Supplementary Information

Supplementary Methods

Necropsy protocols for gastrointestinal tumors

For necropsy of *Pdx-1-Cre*- and *Villin-Cre*-positive mice, the entire gastrointestinal tract was immediately removed, and the stomach was incised along the greater curvature. The intestine was dissected into two pieces and designated as colon and small intestine, which were cut into equal thirds (duodenum, jejunum, and ileum). Each piece of intestine was carefully opened longitudinally with a scissors, and spread onto a piece of filter paper. After flushing with ice-cold saline, the intestine was rolled, luminal side up from its proximal to distal end, onto a wooden stick and the ends of each roll were fixed in place with a 26-gauge needle. After 1 day of fixation in neutral buffered 10% formalin, the stomach was cut into six strips and the rolls of intestine were cut in half with a razor blade along the duodenal to the ileal axis. Then, 5 μm paraffin-embedded sections were stained with hematoxylin and eosin (H&E). Tumor-free intervals were compared by Log-rank test using Prism (GraphPad, La Jolla, CA). When no carcinomatous lesions were identified in a given organ, it was censored for the development of carcinoma on the day of necropsy. Microscopic lung metastases were visualized on a 5 μm paraffin-embedded section containing the largest cross-sectional area of lung.

Immunohistochemistry and immunofluorescence analyses

The excised tumor was fixed in neutral buffered 10% formalin, processed by standard methods and embedded in paraffin. 5-μm cross paraffin sections were dewaxed, rehydrated and subjected to antigen retrieval for immunostaining by heating at 100 °C for 20 minutes in 0.01 M citrate buffer (pH 6.0). The ABC method (Vectastain Elite ABC kit and Vectastain M.O.M. kit, Vector Laboratories, Burlingame, CA) was used according to the manufacturer's protocol. Briefly, the slides were immersed 3% hydrogen peroxide for 10 minutes in order to block endogenouse peroxidase activity. To reduce non-specific binding, the slides were incubated with blocking reagents with the kit. The slides were then incubated for 30 minutes at room temperature with the diluted primary antibodies. The sections were then incubated with biotinylated secondary antibody for 30 minutes at room temperature, followed by incubation with ABC reagent for 30 minutes at room temperature. Subsequently, the slides were subjected to colorimetric detection with ImmPact DAB substrate (SK-4105, Vector Laboratories). The slides were counterstained with Meyer's hematoxylin for 10 seconds. Negative controls were performed by omitting the primary antibody and substitution with diluent. The stain that was unequivocally deeper than the background was identified to be positively stained for each marker. The following antibodies were used in this study; rabbit polyclonal anti E-cadherin

antibody (1:200; Cell Signaling, #3195), rabbit polyclonal anti p53 antibody (1:50; Santa Cruz, sc-6243), mouse monoclonal anti *β*-catenin antibody (1:200; BD, 610154), rabbit monoclonal anti Vimentin antibody (1:100; Cell signaling, #5741), mouse monoclonal anti spasmolytic polypeptide antibody (1:50; Abcam, ab49536), mouse monoclonal anti mucin 6 antibody (1:100; Novus Biologicals, NB120-11335), mouse monoclonal anti mucin 5ac antibody (1:100; Abcam, ab3649), rabbit polyclonal anti somatostatin antibody (1:1000; Immunostar, 20067), rabbit monoclonal anti MMP7 antibody (1:100; Cell signaling, #3801), and rabbit monoclonal anti MMP8 antibody (1:100; Abcam, ab81286). Rabbit polyclonal anti-E-cadherin antibody (1:200; Cell Signaling, #3195, Danvers, MA) was used for both mouse and human tissue samples. For Ki-67 immunofluorescence staining, rabbit polyclonal anti-Ki-67 antibody (1:200; Abcam, ab15580) was used. At least one *200x* microscopic field was examined by a pathologist, to assess the percentage of Ki-67 positive cells. For β-catenin immunostaining, mouse monoclonal anti-*β*-catenin antibody (1:200; BD Transduction Laboratories, 610154, San Diego, CA) was used. Positive cells for *β*-catenin was defined as tumor cells showing a distinct nuclear staining. For vimentin immunostaining, rabbit monoclonal anti vimentin antibody (1:100; Cell signaling Technology, #5741) was used. A positive vimentin staining was defined as a strong cytoplasmic signal equivalent to that of normal mesenchymal tissues. For MMP7 immunostaining, rabbit monoclonal anti MMP7 antibody (1:100; Cell signaling Technology, #3801) was used. A positive MMP7 staining was defined as a strong cytoplasmic signal which was unequivocally deeper than the background.

Stained slides were scanned using ScanScope XT (Aperio, Vista, CA). Scanned images were analyzed at *200x* microscopic fields by a pathologist (J.P). The invasive front of the tumor was defined as the microscopic interface between normal tissue of the host mouse and the tumor mass invading the submucosa or deeper regions of the stomach. Three *200x* microscopic fields at the invasive front were examined by a pathologist to determine the average percentage of cells with nuclear β-catenin staining. The invasive front of the tumor was defined as the microscopic interface between normal tissue of the host mouse and the tumor mass invading submucosa or deeper parts. Unequivocal immunohistochemical staining in tumor nucleus, as compared with normal gastric mucosa, was considered as the positive nuclear β-catenin staining. Differences in the percentage of positive immunostaining between genotypes were evaluated using Student *t*-test.

For E-cadherin immunohistochemistry analyses on clinical samples, 5-μm cryosection slides were obtained from fresh frozen gastric cancer tissue samples, fixed with acetone for 10 minutes, and subject to immunostaining using rabbit polyclonal anti E-cadherin antibody (1:200; Cell Signaling, #3195).

For immunofluorescence analyses, primary antibodies were incubated overnight at 4 °C and secondary antibodies (FITC goat anti-rabbit secondary antibody (1:250; Vector Laboratories FI-1000)) were incubated for 30 minutes at room temperature. Slides were mounted with Vectashield mounting media (Vector Laboratories, H-1200). Mouse monoclonal anti Smad4 antibody (1:50; Santa Cruz, sc-

7966) and rabbit polyclonal anti Ki-67 antibody (1:200; Abcam, ab15580) were used as primary antibodies. Stained slides were scanned using Axiovert 200M (Carl Zeiss, Oberkochen, Germany). Scanned images were analyzed at *200x* microscopic fields by a pathologist (J.P). At least one *200x* microscopic field was examined to assess the percentage of Ki-67 positive cells.

Comparison of gastric adenocarcinoma-free survival

We evaluated the difference in gastric adenocarcinoma-free survival after adjusting for the timing of necropsy using SAS software. By specifying the timing of necropsy as a stratifying variable and by specifying the genotype cohort in the "GROUP=" option, we performed a stratified log-rank test between cohorts. Stratified log-rank test was performed after dichotomizing the timing of necropsy (≤6 *vs* >6 months). The differences in gastric adenocarcinoma-free survival were also calculated with different cutoffs for dichotomizing the timing of necropsy (7 month and 8 month). This p value for stratified log-rank test was presented with standard log-rank test.

Immunoblot analysis

Cells were collected and lysed with T-PER Tissue Protein Extraction Reagent (Thermo Fisher Scientific, Hudson, NH, U.S.A) supplemented with protease inhibitor (0.8 μM aprotinin, 20 μM leupeptin, 10 μM pepstatin A, 40 μM bestatin, and 1 mM phenylmethylsulfonyl fluoride (PMSF)) and phosphatase inhibitor (1 mM sodium fluoride, 1 mM sodium pyrophosphate dehydrate, and 1 mM sodium orthovanadate). After removal of cellular debris, protein concentration was quantitated by using a BCA reagent kit (Thermo Fisher Scientific) according to manufacturer's instruction. Protein sample was prepared by making a 3 in 4 dilution with 4x Laemmli sample buffer (250 mM Tris-HCl (pH 6.8), 4% SDS, 40% glycerol, 0.05% bromphenol blue, 4% 2-mercaptoethanol) and boiling for 5 minute. Equal amounts of protein were separated on SDS-polyacrylamide gel and transferred onto nitrocellulose membrane by electrophoresis and blotting apparatus (Bio-Rad, Hercules, CA). The proteins were probed with the relevant primary antibodies and horseradish peroxidase (HRP)-conjugated secondary antibodies at the recommended dilutions. Rabbit polyclonal anti E-cadherin antibody (1:1000; #3195, Cell Signaling, Danvers, MA), rabbit monoclonal anti acetyl H3K9 (1:1000; #9649, Cell Signaling Technology), mouse monoclonal anti *β*-catenin antibody (1:1000; 610154, BD Transduction Laboratories), and mouse monoclonal anti GAPDH (1:1000; sc-32233, Santa Cruz, CA) were applied. Immunodetection were performed by using an enhanced chemiluminescence (ECL) detection kit (Thermo Fisher Scientific).

Genotyping PCR

Mouse tail genomic DNA was isolated using Genomic DNA Mini Kit (Geneaid, New Taipei, Taiwan). PCR genotyping primers for Cdh1 were F: 5'-CTTATACCGCTCGAGAGCCGGA-3' and R: 5'-GTGTCCCTCCAAATCCGATA-3' [1]. Amplicons of 900 and 980 bp were expected for wild-type and

floxed alleles, respectively. PCR genotyping primers for *Trp53* were F: 5'-TGGAGATATGGCTTGGAGTAG-3' and R: 5'-CAACTTACTTCGAGGCTTGTC-3'. PCR products of 420 and 500 bp were expected for wild-type and floxed alleles, respectively. PCR genotyping primers for *Smad4* were F: 5'-GGGCAGCGTAGCATATAAGA-3' and R: 5'-GACCCAAACGTCACCTTCAC-3'. PCR products of 390 and 480 bp were expected for wild-type and floxed alleles, respectively. Primers for *Pdx1-Cre* were F: 5'-CTGGACTACATCTTGAGTTGC-3' and R: 5'-CAGATTACGTATATCCTGGCAG-3'. Primers for *Villin-Cre* were F: 5'-TCCTCTAGGCTCGTCCCG-3' and R: 5'-CAGATTACGTATATCCTGGCAG-3'. Primers for *MMTV-Cre* were F: 5'-GTCGATGCAACGAGTGATGAG-3' and R: 5'-TCATCAGCTACACCAGAGACG-3'.

Quantitative real-time RT-PCR (QRT-PCR)

For mouse samples, total RNA was isolated using AllPrep DNA/RNA/Protein Mini Kit (Qiagen, Valencia, CA) according to the manufacturer's instructions. For RNA isolation from fresh frozen tissues, stroma was trimmed out using a H&E-stained top slide, before RNA isolation. For RNA isolation from primary cultured cells, 1.5 x 10^5 cells were seeded into 12 wells and harvested at 24 hours after seeding. To treat the primary cells with BMP and TGF ligands, 1.5 x 10^5 cells were equally seeded into 12 wells and each well was treated with 100ng/ml BMP-2 (355-BM, R&D Systems, Minneapolis, MN), 400 ng/ml of noggin (6057-NG, R&D Systems, Minneapolis, MN), 5 ng/ml of TGF-β1 (gift from Dr. Lalage Wakefield) at 24 hours after seeding. 16 hours after the treatment, these cells were harvested. 0.3 μg of total RNA was reverse transcribed using random hexamers and *amfiRivertII* Reverse Transcriptase (GenDEPOT, Barker, TX) according to the manufacturer's standard protocols. PCR reactions were performed on a Roche LC480 (Roche Diagnostics, Penzberg, Germany) using 5 μl of 2× QuantiTect SYBR Green PCR Master Mix (Qiagen), 400 nM of each primer, and 2 μl of cDNA sample which was diluted 1:5 in water in a total volume of 10 μl. Cycling conditions were as follows: 15 min at 95ºC, followed by 55 cycles each consisting of 20 s at 94ºC, 20 s at 57ºC and 20 s at 72ºC. Data were analyzed using the LC480 software (Roche Diagnostics). QRT-PCR primers were F: 5'-GCTGCAGGTCTCCTCATG-3' and R: 5'-CATCCTTCAAATCTCACTCTGC-3' for *Cdh1,* F: 5'-ACGATGAGGACCAGGTGGTAG-3' and R: 5'-CAGTACAACGAGCTGTCTCTAC-3' for *Ctnnb1,* F: 5'-GTGCGCCAGCAGTATGAAAG-3' and R: 5'-GCATCGTTGTTCCGGTTGG-3' for *Vim,* F: 5'-ACTTACCTCGGATCGTAGTG-3' and R: 5'-TCTCCATGATCTCTCCTTGC-3' for *Mmp7,* F: 5'-ATCCTGGTGCCTTGATGTAC-3' and R: 5'-GTCTGAAGGTCCATAGATTGTC-3' for *Mmp8,* F: 5'-CCAGCAGATTTCAAGGTGGAC-3' and R: 5'-TTACAGCTACCTGCCACTTTTC-3' for *Cdh2,* F: 5'-CACACGCTGCCTTGTGTCT-3' and R: 5'-GGTCAGCAAAAGCACGGTT-3' for *Snai1,* F: 5'-CCTTGGGGCGTGTAAGTCC-3' and R: 5'-TTCTCAGCTTCGATGGCATGG-3' for *Snai2,* F: 5'-TGATGAAAACGGAACACCAGATG-3' and R: 5'-GTTGTCCTCGTTCTTCTCATGG-3' for *Zeb1,* F: 5'-AGCGACACGGCCATTATTTAC-3' and R: 5'-GTTGGGCAAAAGCATCTGGAG-3' for *Zeb2,* F: 5'-

GGACAAGCTGAGCAAGATTCA-3' and R: 5'-CGGAGAAGGCGTAGCTGAG-3' for *Twist1, ,* F: 5'-ACGAGCGTCTCAGCTACGCC-3' and R: 5'-AGGTGGGTCCTGGCTTGCGG-3' for *Twist2* and F: 5'-GAACATGGCATTGTTACCAACTG-3' and R: 5'-GTGTTGAAGGTCTCAAACATGATC -3' for *ActB.* The $2^{-\Delta\Delta CT}$ method was used to calculate relative changes [2].

For human samples, total RNA was isolated from fresh frozen tissue using *mir*Vana™ Isolation Kit (Ambion), according to the manufacturer's instructions. Stroma was trimmed out using a H&E-stained top slide, before RNA isolation. Isolated RNA was treated with DNase I (Sigma, St. Louis, MO) according to the manufacturer's protocol. 0.5 µg RNA was reverse transcribed using random hexamers by using SuperScript III First-Strand Synthesis System (Invitrogen) according to the manufacturer's standard protocol.

QRT-PCR primers were F: 5'-CGCATTGCCACATACACTCTC-3' and R: 5'-GGTTCCTGGAAGAGCACCTTC-3' for *CDH1*, and F: 5'-GAGTCAACGGATTTGGTCG-3' R: 5'-TGGAATCATATTGGAACATGTAAAC-3' for *GAPDH1*.

Genomic DNA quantitative real-time PCR (QPCR)

Genomic DNA was isolated from 5 fresh frozen tissue samples and 13 formalin-fixed paraffin-embedded blocks of gastric cancers formed in *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* mice, using AllPrep DNA/RNA/Protein Mini Kit (Qiagen) and QIAamp DNA FFPE Tissue Kit (Qiagen) and respectively according to the manufacturer's instructions. Using macrodissection, stromal tissue was trimmed out as much as possible, to enrich tumor cells. Genomic DNA was subject to the RNAse I treatment at RT for 2 min according to the manufacturer's standard protocol (Qiagen). Concentration of DNA was measured using NanoDrop (Thermo Fisher Scientific, Hudson, NH). PCR reactions were performed in a total volume of 10 µl containing 5 µl of QuantiTect Probe PCR kit (Qiagen), 400 nM of each primer, 200 nM of TaqMan probe and 60 ng of genomic DNA using a Roche LC480 (Roche Diagnostics). PCR conditions were as follows: 15 min at 95ºC, followed by 99 cycles each consisting of 20 s at 94ºC, 20 s at 55ºC, and 20 s at 72ºC. Data were analyzed using the LC480 software (Roche Diagnostics). TaqMan probe/primers were targeted to exons 6, 8, and 10 of the mouse *Cdh1* gene. The QPCR primers were as follows: Cdh1-Exon6-F/R, 5'-TCTATTCTCATGCCGTGTCATC-3'/5'-GCCTGTTGTCATTCTGATCTGTC-3'; Cdh1-Exon8-F/R, 5'-CCTACATACACTCTGGTGGTTC-3'/5'-CGTGCTTGGGTTGAAGACAG-3'; Cdh1-Exon10-F/R, 5'-AGCAGCAATACATCCTTCATG-3'/5'-GTCTACCACGTCCACAGTGAC-5', specifically for exon6, exon8 and exon10 (between LoxP luci) of *Cdh1* focusing on only non-functional *Cdh1* allele; Rnu6-F/R, 5'-GCTTCGGCAGCACATATACTA-3'/5'-TTGCGTGTCATCCTTGCGCAG-3'. The TaqMan probes were as follows: Cdh1-Exon6-Probe, 5'-FAM-AGGATCCCATGGAGATAGTGATCACAGT-BHQ-1-3'; Cdh1-Exon8-Probe, 5'-FAM-CTGACCTTCAAGGTGAAGGCTTGAGC-BHQ-1-3'; Cdh1-Exon10-Probe, 5'-FAM-TCTCTTGTCCCTTCCACAGCCACT-BHQ-1-3'; Rnu6-Probe, 5'-FAM-ATTGGAACGATACAGAGAAGATTAGCATGG-BHQ-1-3'.

Rnu6 gene was used as a loading control. Tumor data were normalized against data obtained for genomic DNA obtained from *Pdx-1-Cre*-negative mice, which was serially diluted as standards for quantitation. LOH was defined as the average $\log_2$ ratio of three probes of tumor to normal DNA < -1.5 [3], given the median tumor nuclei content of 60% in gastric cancers that formed in *Pdx-1-Cre; Trp53$^{F/F}$;Cdh1$^{F/+}$* mice.

Expression array analyses

Total RNA was isolated from tumor-rich area of fresh frozen gastric adenocarcinomas formed in 4 *Pdx-1-Cre; Trp53$^{F/F}$;Cdh1$^{F/+}$* mice and from normal gastric mucosa samples obtained from 4 *Cre*-negative mice. 1 µg of total RNA was subjected to GeneChip Mouse Gene 1.0 ST Arrays (Affymetrix, Santa Clara, CA) and summarized with robust multichip average (RMA) using R (version 2.15.2). Student *t*-test was used to identify differentially expressed genes. Pathway analysis was performed using Gene Set Enrichment Analysis (GSEA) (http://www.broadinstitute.org/gsea). Genes differentially expressed between tumor and normal samples at P<0.001 were subject to the GSEA of canonical pathways and transcription factor targets.

Affymetrix's GeneChip miRNA 2.0 Array was used for microRNA analysis. Microarray data were normalized using R and Student t-test was used to identify microRNAs differentially expressed between *Pdx-1-Cre;Trp53$^{F/F}$;Cdh1$^{F/+}$* gastric adenocarcinomas (n=2) and normal tissue (n=2). A P value less than 0.001 was used as the cutoff to identify differentially expressed microRNAs.

Methylation Specific PCR

For mouse samples, genomic DNA was isolated from gastric adenocarcinomas that formed in *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* mice using AllPrep DNA/RNA/Protein Mini Kit (Qiagen) according to the manufacturer's instructions. Using macrodissection, stromal tissue was trimmed out as much as possible. Each 0.3 µg of genomic DNA was bisulfite-converted by using EZ DNA Methylation^{TM} kit (Zymo Research, Irvine, CA) according to the manufacturer's protocol. Methylation-specific PCR was performed in a reaction volume of 10 µl consisting of 0.05 µl of HotStarTaq (Qiagen), 0.8 µl of 2.5 mM dNTP, 1 µl of 10× PCR buffer, 2 µl of 5× Q-Solution, 400 nM of each primer and 1.5 µl of bisulfite-modified (20 µl-eluted) genomic DNA sample, with an initial activation at 95ºC for 15 min, followed by 35 cycles (94ºC for 45 s, 55ºC for 35 s and 72ºC for 40 s) and a final extension at 72ºC for 10 min. The following primers were used- F: 5'-TGTTTATTGGTGTGGGAGTTGTG-3' and R: 5'-CAAAACCCTCCACATACCTACAAC-3' to detect unmethylated CpG sites, and F: 5'-GTTTATTGGTGTGGGAGTCGC-3' and R: 5'-CAAAACCCTCCACATACCTACAAC-3' for methylated CpG sites.

For human samples, genomic DNA samples were isolated from fresh frozen gastric tissue of 3 E-cadherin-negative patients, using DNeasy Blood & Tissue Kit (Qiagen) according to the manufacturer's protocol. Methylation-specific PCR for human *CDH1* was performed as described

above. Primers were designed as reported previously [4]. The following primers were targeted to the CpG island 1; F: 5'-TAATTTTAGCTTAGAGGGTTATTGT-3' and R: 5'-CACAACCAATCAACAACACA-3' for unmethylated CpG sites, and F: 5'-TTAGGTTAGAGGGTTATCGCGT-3' and R: 5'-TAACTAAAAATTCACCTACCGAC-3' for methylated CpG sites. The following primers were targeted to the CpG island 2; F: 5'-GGTGGGTGGGTTGTTAGTTTTGT-3' and R: 5'-AACTCACAAATCTTTACAATTCCAACA-3' for unmethylated CpG sites, and F: 5'-GTGGGCGGGTCGTTAGTTTC-3' and R: 5'-CTCACAAATACTTTACAATTCCGACG-3' for methylated CpG sites. The following primers were targeted to the CpG island 3; F: 5'-GGTAGGTGAATTTTTAGTTAATTAGTGGTA-3' and R: 5'-ACCCATAACTAACCAAAAACACCA-3' for unmethylated CpG sites, and F: 5'-GGTGAATTTTTAGTTAATTAGCGGTAC-3' and R: 5'-CATAACTAACCGAAAACGCCG-3' for methylated CpG sites. The following primers were targeted to the CpG island 4; F: 5'-GGGGTGTTTGGTTGTGGAGTTT-3' and R: 5'-TTCCCTCAAAAATCATCCCCAC-3' for unmethylated CpG sites, and F: 5'-GCGTTTGGTCGCGGAGTTC-3' and R: 5'-TTCCCTCAAAAATCGTCCCCAC-3' for methylated CpG sites.

Bisulfite sequencing

Genomic DNA was bisulfite-modified and PCR-amplified with Cdh1-BGS-F (5'-GTGGAATAGGAAGTTGGGAAGTT-3') and Cdh1-MSP-Un-R primers, as in Methylation Specific PCR. After PCR products were purified using gel extraction kit (Macrogen, Seoul, Korea), the purified PCR products were cloned into the TOPO-TA vector (Invitrogen) and transformed of *E-coli* according to the manufacturer's instructions. After isolation of plasmid DNA from 8-10 clones per each sample using plasmid Mini-Prep kit (Macrogen), each plasmid sample was sequenced with M13-F (-20) primer (5'-GTAAAACGACGGCCAG-3').

In vivo 5-aza-2'-deoxycitidine (5-Aza) challenge

5-aza-2'-deoxycytidine (5-Aza) was purchased from Sigma-Aldrich (St. Louis, MO). 5-Aza was dissolved in DMSO to a concentration of 10 mM. Two *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* mice were given 2.5 mg/kg of 5-Aza by intraperitoneal injection in a volume of 1.1 μl DMSO per gram body weight (Day 1). The same dose was repeated on Day 3. Three days after the first 5-Aza dose, mice were sacrificed and fresh frozen gastric cancer tissue was subjected to RNA isolation and real-time RT-PCR as described above.

Trichostatin A treatment

Trichostatin A (TSA) was purchased from Sigma-Aldrich. Gastric cancer of a *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* mouse was harvested at the necropsy, and immediately minced into small tissue fragments using scissors and blades in RPMI-1640 media supplemented with 10% FBS.

Tissue fragments were incubated in the media containing TSA for 24 hours. Then, tissue fragments were collected and total proteins extraction were performed using T-PER Tissue Protein Extraction Reagent (Thermo Fisher Scientific).

Exome sequencing of human samples

Frozen gastric cancer tissue samples and adjacent normal tissue samples were obtained, with the IRB approval and signed informed consents, from 13 gastric cancer patients who underwent gastrectomy at Keimyung University Dongsan Hospital in Taegu, Korea, and at Asan Medical Center in Seoul, Korea (Table 1). Tissue samples were cryosectioned and H&E-stained. Guided by the H&E-stained top slide, macrodissection was performed for each sample. Median tumor nuclei content in all tumor samples was 75%. 20 mg of each of macrodissected normal and cancerous tissues were pulverized using metal mortar and pestle. Tissue powder was subjected to total RNA and genomic DNA isolation, using *mir*Vana™ Kit (Ambion, TX) and DNeasy Blood & Tissue Kit (Qiagen), respectively. Total RNA was treated with DNase I (Sigma) using the following reagents: 1/10 volume of DNase I buffer, 1/10 volume of DNase I, 1/20 volume of dithiothreotol (DTT, Qiagen) and 1/40 volume of RNass inhibitor (RNasin, Promega, WI). Then, the total RNA was purified by using Trizol (Ambion), following manufacturer's instruction. In-solution RNase A (100mg/ml, Qiagen) treatment was performed for isolated genomic DNA. Exome sequencing was performed using HiSeq 2000 instrument (Illumina, Hayward, CA), after capturing using SureSelect Human All Exon V4+UTR kit (Agilent Technologies, Santa Clara, CA). Fastq files were aligned and duplicates were removed using BWA and PICARD, respectively. Somatic single nucleotide variations were identified and annotated using VarScan and Annovar, respectively.

Primary tissue culture and *stable expression of Smad4 in primary cultured cells using lentiviral transduction system*

Tumors were minced into small pieces and digested with collagenase for 30 min at 37˚C. Single tumor cells were seeded into plated on a 60-mm dish in RPMI 1640 media with 20 % fetal bovine serum (FBS), 100 U/ml of penicillin, and 0.1 mg/ml of streptomycin in humidified incubator at 37˚C in 5 % CO_2. For better expression of the transgene in our mouse gastric cancer cell, *Smad4* gene was cloned into a modified lentiviral vector, named pCDH-CAG-MCS-EF1-Puro, from an original vector, pCDH-CMV-MCS-EF1-Puro (System Biosciences). CMV (cytomegalovirus) promoter in the original vector was replaced with CAG (CMV early enhancer/chicken β actin) promoter. Briefly, a CAG promoter cassette, obtained from PCR amplification of the CAG promoter of Ai9 vector (Addgene) with CAG-F (5'-ATACTAGTTATTAATAGTAATCAATTACGGG-3', a *Spe*I site is underlined) and CAG-R (5'-ATGAATTCGCTAGCTTTGCCAAAATGATGAGACAG-3', a *Eco*RI site is underlined) primers, was cloned into the original vector cut with *Spe*I and *Eco*RI. Next, briefly, the gene cassette of mouse *Smad4*, PCR-amplified from cDNA of the mouse cells with Smad4-F (5'-

ATGAATTCGCCACCATGGACAATATGTCTATAACAAATACAC-3', a *Eco*RI site is underlined) and Smad4-R (5'-ATGCGGCCGCTCAGTCTAAAGGCTGTGGGTC-3', a *Not*I site is underlined) primers, was cloned into the modified vector cut with *Eco*RI and *Not*I. All cloned genes were validated by using Sanger sequencing. Lentivirus was produced by cotransfecting Smad4-expressing lentiviral vector and pMD2.G and psPAX2 constructs (Addgene) into 293T cells by using lipofectamine 2000 (Invitrogen). Viral supernatants were harvested 48 hours after transfection, filtered though a 0.45 μm filter, titered, and used to infect primary cultured cancer cells with 10 μg/mL polybrene. Cells were treated by 2 μg/mL puromycin at 48 hours after viral transduction and were selected for 3 days.

Establishment of stable β-catenin knock down cells using lentiviral shRNA and measurement of growth rates

The lentiviral Ctnnb1 shRNA constructs were purchased from Sigma-Aldrich (St. Louis, MO) with pLKO.1-puro eGFP control vector (Sigma, SHC005). The target set was generated from accession number NM_007614.2: (1) CCGGGCGTTATCAAACCCTAGCCTTCTCGAGAAGGCTAGGGTTTGATAACGCTTTTT(2) CCGGCCCAAGCCTTAGTAAACATAACTCGAGTTATGTTTACTAAGGCTTGGGTTTTT. Lentiviruses were produced by cotransfecting shRNA-expressing vector and pMD2.G and psPAX2 constructs (Addgene) into 293T cells by using lipofectamine 2000 (Invitrogen). Viral supernatants were harvested 48 hours after transfection, filtered though a 0.45 μm filter, titered, and used to infect primary cultured cancer cells with 10 μg/mL polybrene. Cells were treated by 2 μg/mL puromycin at 48 hours after viral transduction and were selected for 3 days.

To measure monolayer growth rates of these cells, the cells were seeded on tissue culture 100 mm^2 dishes at 1.0 x 10^5 cells/well in RPMI 1640 media containing 10% FBS. At 24 and 48 hours after seedingcell proliferation was determined by manual cell counting using the trypan blue exclusion assay.

Luciferase reporter assay

β-catenin activity and TGFβ activity was evaluated by using Cignal TCF/LEF reporter assay kit and SMAD reporter assay kit (CCS-018L and CCS-017L, SA Biosciences, Frederick, MD). To measure BMP activity, primary cultured cells were cotransfected with a well-characterized BMP-response element (250 ng/ml, BRE-Luc, gift from Dr. Tae-Aug Kim) and Renilla plasmid (250 ng/ml) [5]. 1.5 x 10^5 primay cultured cells were suspended in 1ml of Opti-MEM medium (Life Technologies, Grand Island, NY) and the suspension cells were seeded in 12 well plates. These cells were transiently transfected in suspension with the Tcf/Lef reporter plasmid, Smad reporter plasmid, and BRE-Luc reporter plasmid using Lipofectamine 2000 transfection reagent (Life Technologies). 24 hours after transfection, Opti-MEM medium was changed to RPMI 1640 containing 0.5% FBS with 100 ng/mL of BMP2 (355-Bm, R&D Systems, Minneapolis, MN) or 400 ng/mL of Noggin (6057-NG, R&D Systems,

Minneapolis, MN) or 5 ng/ml of TGF-β1 (gift from Dr. Lalage Wakefield). 16 hours after the treatment, luciferase assays were carried out using the dual luciferase reporter assay system (Promega, Madison, WI) according to the manufacturer's protocol. Wnt3a conditioned medium produced by L Wnt-3A cells (gift from Dr. Tae-Aug Kim) was also used as a positive control for Tcf/Lef reporter activity. Light emission was quantified with a Victor 3 1420 luminescence microplate reader (Perkin-Elmer, Waltham, MA). The signals were normalized for transfection efficiency to the internal *Renilla* control.

PCR test for Helicobacter pylori (H. pylori)

Genomic DNA was isolated from fresh frozen tissue sample obtained from adjacent normal tissue of 13 young gastric cancer patients analyzed in this study. Primers to detect the 16S rDNA gene of *H. pylori* were designed as described in a study of *H. pylori* murine vaccine model [5]. PCR reaction mixture contained 0.15 μl of *i*-StarTaq (Intron), 1.5 μl of 2.5 mM dNTP, 1.5 μl of 10× PCR buffer, 500 nM of each primer (F:5'-TTTGTTAGAGAAGATAATGACGGTATCTAAC-3' and R:5'-CATAGGATTTCACACCTGACTGACTATC-3'), 3 μl of 5× Q-solution (Qiagen) and 30 ng of genomic DNA in a final volume of 15 μl. Cycling conditions were as follows: 2 min at 95ºC, followed by 60 cycles each consisting of 35 s at 94ºC, 30 s at 57ºC and 30 s at 72ºC, and 3 min at 72ºC. Distilled water and *H. pylori* DNA (1 mM and 0.1 mM, mixed with SNU-16 cells), was used as negative and positive controls, respectively.

Migration assay

Primary cultured cells were plated on 24-well inserts with an 8 μm pore size (353097, BD) at 2.5 x 10^4 cells/well in serum-free RPMI 1640 media. Media containing 10% FBS was added to the 24-well lower insert chambers (354578, BD) and cells were cultured at 37°C and 5% CO_2. 24 h after plating, cells remaining in the upper chamber were removed by gently scrapping the upper chamber with a wet cotton swab. Cells that had migrated to the lower insert chamber were fixed with 10% formalin for 10 min and washed one with PBS. The inserts were soaked in hematoxylin for 1 min and eosin for 5 min and then washed with tap water several times. Membranes were cut from the inserts and placed on a glass slide and mounted under a coverslip with Permount (Fisher Scientific, Waltham, MA). The number of migrated cells was averaged by counting 3 high power fields (*200x*) and compared between groups using Student *t*-test.

References

1. Boussadia O, Kutsch S, Hierholzer A, et al. E-cadherin is a survival factor the lactating mouse mammary gland. Mech Develop 2002;115:53-62.

2. Livak MJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-CT) method. Methods 2001;25:402-408.

3. Gutierrez A, Dahlberg SE, Neuberg DS, Zhang J, Grebliunaite R, Sanda T, et al. Absence of biallelic TCRγ deletion predicts early treatment failure in pediatric T-cell acute lymphoblastic leukemia. J Clin Oncol 2010 20;28(24):3816-23.

4. Graff JR, Herman JG, Myöhänen S, et al. Mapping Patterns of CpG Island Methylation in Normal and Neoplastic Cells Implicates Both Upstream and Downstream Regions in *de Novo* Methylation. J Biol Chem 1997;35:22322-22329.

5. Korchynskyi O, ten Dijke P. Identification and functional characterization of distinct critically important bone morphogenetic protein-specific response elements in the Id1 promoter. *J Biol Chem* 2002;277:4883–4891.

6. Roussel Y, Harris A, Lee MH, et al. Novel methods of quantitative real-time PCR data analysis in a murine *Helicobacter pylori* vaccine model. Vaccine 2007;25:2919-2929.

Supplementary Result

Comparison of gastric adenocarcinoma-free survival

The difference in gastric adenocarcinoma-free survival between *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* and *Pdx-1-Cre;Trp53$^{F/F}$;Cdh1$^{F/+}$* cohorts was statistically significant for the stratified log-rank test (p=0.007), when the timing of necropsy was dichotomized as ≤6 months *vs* >6 months. The differences in gastric adenocarcinoma-free survival remained significant when different cutoffs were used for dichotomizing the timing of necropsy [p=0.003 and p=0.019 for 7 month and 8 month, respectively]. The difference in gastric adenocarcinoma-free survival between *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* and *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{+/+}$* cohorts was p<0.001 for the stratified log-rank test throughout different cutoffs for dichotomizing the timing of necropsy (6 month, 7 month and 8 month).

Supplementary Tables

Table S1. Incidence of adenocarcinomas in each organ according to the genotype

Genotype	Median gastric adenocarcinoma-free survival	Stomach	Duodenum	Pancreas	Distant Metastasis
Pdx1Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$	8.0 mo	21/25 (84%)	9/25 (36%)	2/25 (8%)	3/25 (12%)
Pdx1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$	Not reached	1/28 (0%)	13/28 (46%)	1/28 (3.6%)	0/28 (0%)
Pdx1-Cre;Trp53$^{F/F}$;Cdh1$^{F/F}$	9.4 mo	6/15 (40%)	0/15 (0%)	1/15 (6.7%)	0/15 (0%)
Pdx1-Cre;Trp53$^{F/F}$;Cdh1$^{F/+}$	Not reached	0/7 (0%)	0/7 (0)	0/7 (0%)	0/7 (0%)

Table S2. Gene sets most enriched in 1,096 genes upregulated in gastric adenocarcinomas at $P < 0.001$

Gene set	**Transcription Factor**	***P***	***No.* genes in gene set**	***No.* overlapping Genes**
CAGGTG_V$E12_Q6	*Tcf3*	$<10^{-16}$	2485	162
CAGCTG_V$AP4_Q5	*Repin1*	$<10^{-16}$	1524	106
GATTGGY_V$NFY_Q6_01		$<10^{-16}$	1160	81
TTGTTT_V$FOXO4_01	*Mllt7*	$<10^{-16}$	2061	151
CTTTGA_V$LEF1_Q2	*Lef1*	$<10^{-16}$	1232	92
TATAAA_V$TATA_01	*Taf*	$<10^{-16}$	1296	100
GGGCGGR_V$SP1_Q6	*Sp1*	$<10^{-16}$	2940	229
GGGAGGRR_V$MAZ_Q6	*Maz*	$<10^{-16}$	2274	178
TGACAGNY_V$MEIS1_01	*Meis1*	$<10^{-16}$	827	65
GGGTGGRR_V$PAX4_03	*Pax4*	$<10^{-16}$	1294	102
AACTTT_UNKNOWN		$<10^{-16}$	1890	151
CTGCAGY_UNKNOWN		$<10^{-16}$	765	65
RNGTGGGC_UNKNOWN		$<10^{-16}$	766	66
TGGAAA_V$NFAT_Q4_01	*Nfat*	$<10^{-16}$	1896	164
TGANTCA_V$AP1_C	*Jun*	$<10^{-16}$	1121	99
CTTTGT_V$LEF1_Q2	*Lef1*	$<10^{-16}$	1972	182
RYTTCCTG_V$ETS2_B	*Ets2*	$<10^{-16}$	1085	111
TTTNNANAGCYR_UNKNOWN		$<10^{-16}$	133	29
KRCTCNNNNMANAGC_UNKNOWN		$<10^{-16}$	66	25
CACGTG_V$MYC_Q2	*v-Myc*	1.1×10^{-16}	1032	73

Table S3. Clinicopathologic characteristics of young gastric cancer patients

	E-cadherin immunostaining	
	Positive *(n=10)*	**Negative *(n=3)***
Median age (yr)	35.5	36.5
Female	6 (60%)	1 (33%)
***H. pylori*-negative**	2 (20%)	2 (66.7%)
Histologic type		
Poorly differentiated tubular	5 (50%)	2 (66.7%)
Signet ring cell carcinoma	5 (50%)	1 (33.3%)
Stage, AJCC		
II	1 (10%)	2 (66.7%)
III	2 (20%)	1 (33.3%)
IV	7 (70%)	0

AJCC, American Joint Committee on Cancer (7^{th} Edition)

Supplementary Figure Legends

Fig S1. Timing of necropsy was not statistically different across genotypes, P value for ANOVA=0.33.

Fig S2. Immunofluorescence and immunohistochemistry for the confirmation of loss of Smad4 and p53 in *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* tumors.

Fig S3. Immunohistochemistry for the tissue origin of the gastric adenocarcinomas arising in *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* mice.

Fig S4. Histological findings of mammary adenocarcinoma arising in *MMTV-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$. (A)* These tumors showed morphological features of invasive ductal adenocarcinoma (B) with a squamous component.

Fig S5. Immunohistochemistry for the tumor center and invasive front of gastric cancers formed in *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* mice

Fig S6. MMP8 expression at the invasive front of gastric cancer of *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* mice as compared with that of *Pdx-1-Cre;Trp53$^{F/F}$;Cdh1$^{F/+}$* mice. (A) Representative MMP8 immunohistochemical staining across the genotypes. (B) Percentage of MMP8-positive cells in gastric cancer tissues arising in *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* mice (n=10) and *Pdx-1-Cre;Trp53$^{F/F}$;Cdh1$^{F/F}$* mice (n=5).

Fig S7. (A) Activity of BMP-specific reporter (BRE-Luc) in cells used for Fig 7A and B. Reporter activity was measured after treatment with 100 ng/ml of BMP2, 400 ng/ml of Noggin, and 5 ng/ml of TGF-β1 for 16 hours. BMP2 treatment in Smad4 expressing cells led to specific increase in BRE-Luc reporter activity. (B) TGF-β reporter activity of cells used for Fig 7A and B after treatment with 5 ng/ml TGF-β1 and 100 ng/ml of BMP2 for 16 hours. TGF-β1 treatment in Smad4 expressing cells led to specific increase in TGF-β reporter activity.

Fig S8. (A) mRNA expression of *MMP7, MMP8* and EMT-activating transcription factors of *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* gastric cancer cell lines (n=4) as compared with that of *Pdx-1-Cre;Trp53$^{F/F}$;Cdh1$^{F/F}$* gastric cancer cell lines (n=2). (B) Western blot analyses showed vimentin expression in primary cell lines from gastric cancer tissues arising in *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* mice.

Fig S9. Monolayer growth rate of primary cells after knockdown of β-catenin in the *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* cell lines. There were no differences in monolayer growth at Day 1 (when migration assay was performed) and Day 2 between β-catenin knockdown cells and control cells.

Fig S10. (A) Western blot analyses on two *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* cell lines (used for migration assays) showed unchanged, almost undetectable E-cahderin expression after knockdown of β-catenin. (B) Genomic DNA real-time PCR for the *Cdh1* gene confirmed that these *Pdx-1-Cre;Smad4$^{F/F}$;Trp53$^{F/F}$;Cdh1$^{F/+}$* cell lines retained one wild-type *Cdh1* allele ($\log_2$ ratio= -1), suggesting that loss of E-cadherin protein of these cells is not due to LOH.