Regulation of TGF-β signaling and EMT in cancer progression
Zhang, J.

Chapter 3

Opposing USP19 splice variants in TGF-β-induced breast cancer cell epithelial-mesenchymal transition

Jing Zhang[1,2], Maarten van Dinther[1,2], Midory Thorikay[1,2], Babak Mousavi Gourabi[3], Boudewijn PT Kruithof[2,4], Peter ten Dijke[1,2]*

[1]Oncode Institute, Leiden University Medical Center, 2300 RC Leiden, The Netherlands.

[2]Cell Chemical Biology, Leiden University Medical Center, 2300 RC Leiden, The Netherlands.

[3]Dept. of Anatomy and Embryology, Leiden University Medical Center, 2300 RC Leiden, The Netherlands.

[4]HARTZ, Leiden University Medical Center, 2300 RC Leiden, The Netherlands.

Chapter 3

Abstract

The ubiquitin specific protease (USP)19 is a deubiquitinating enzyme that regulates the stability and function of multiple proteins thereby controlling various biological responses. Alternative splicing of USP19 results in two major encoded variants that are localized in endoplasmic reticulum (ER) (USP19-ER) and cytoplasm (USP19-CY). However, the importance of alternative splicing for USP19 function remains unclear. Here we demonstrated that the two USP19 isoforms regulate transforming growth factor (TGF)-β/SMAD signaling and subsequent biological responses in an opposing manner in breast and lung cancer cells. The USP19-CY isoform promotes the TGF-β-induced signals by directly interacting with TGF-β type I receptor (TβRI), inhibiting its polyubiquitination, thereby stabilizing TβRI levels at the plasma membrane. Oppositely, USP19-ER binds and sequesters TβRI in the ER. By decreasing cell surface TβRI levels, USP19-ER inhibits TGF-β/SMAD signaling in a deubiquitination independent manner. Moreover, USP19-ER inhibits TGF-β-induced EMT of breast and lung cancer cells, whereas USP19-CY enhances EMT in a deubiquitinase dependent manner, as well as migration and extravasation of cancer cells in the zebrafish xenograft model. Furthermore, we observed that USP19-CY is highly expressed in breast cancer tissues compared to adjacent normal tissues, suggesting that its expression is correlated with poor prognosis. Notably, small molecule splicing modulator herboxidiene inhibited USP19-CY and promoted USP19-ER expression and inhibited breast cancer cell migration. Such splicing modulators or compounds that inhibit USP19 deubiquitinating activity may have therapeutic potential for breast and lung cancer.

Introduction

Transforming growth factor-β (TGF-β) is a multifunctional cytokine, which fulfils essential roles in the development and homeostasis in most human tissues [1, 2]. Perturbation of TGF-β signaling has been linked to multitude of human diseases, including cancer [2, 3]. The initiation of TGF-β signaling starts from the binding of TGF-β to the extracellular domain of the transmembrane TGF-β type II receptor (TβRII) that is

endowed with serine/threonine kinase activity [4, 5]. Then the TGF-β type I receptor (TβRI) is recruited to form a heteromeric complex with TβRII and becomes trans-phosphorylated by the TβRII kinase [6]. Activation of the TβRII/TβRI complex phosphorylates the Sma- and Mad-related (SMAD) proteins, i.e., SMAD2 and SMAD3, at two carboxy-terminal serine residues. These phosphorylated R-SMADs form complexes with a common SMAD mediator, i.e., SMAD4, and translocate into the nucleus to interact with high-affinity DNA binding transcription factors and chromatin remodelling proteins, thereby modulating the transcription of TGF-β target genes including *SMAD7*, *SERPINE1*, *CCN2*, encoding the protein SMAD7, PAI1, and CTGF, respectively [7, 8]. TGF-β is a strong driver of epithelial-mesenchymal transition (EMT), which is a dynamic and reversible process that epithelial cells lose their cell–cell contacts and apical–basal polarity, and gain mesenchymal phenotypes with enhanced migratory abilities [9]. EMT is characterized by the downregulation of epithelial markers, i.e., E-cadherin and claudin-1, and upregulation of mesenchymal markers, i.e., N-cadherin, vimentin, and SNAIL1/2 [10]. The incomplete epithelial to mesenchymal transition is referred to as epithelial-mesenchymal plasticity (EMP) [11]. EMT plays a critical role in embryonic development [12] and cancer cell migration, invasion and metastasis [13-15].

Ubiquitination of TGF-β signaling components, including its receptors is a post-translational modification that is emerging as a key mechanism by which TGF-β signaling is kept in check [16, 17]. Ubiquitination depends on a cascade of enzymes consisting of ubiquitin-activating enzymes (E1), ubiquitin-conjugating enzymes (E2), and ubiquitin ligases (E3) to mediate the transfer of ubiquitin to the protein substrates [18]. The TGF-β target gene *SMAD7*, a negatively regulator of TGF-β signaling, binds to SMURF E3 ubiquitin ligases and brings it to the TβRI, leading to SMURF-mediated polyubiquitylation routes of the receptor for degradative endocytosis [19-21]. Directly opposing the conjugating function of E3 ligases are the deubiquitinases (DUBs), which remove the ubiquitin chains from target proteins [22]. The DUBs, ubiquitin specific protease (USP)4 and USP15 have been shown to antagonize SMAD7/SMRF2 mediated poly-ubiquitination and proteasomal degradation of TβRI. Whereas USP4

3

was found to directly interact with TβRI [23], USP15 was recruited by SMAD7 to TβRI [24, 25]. In addition, USP4 was found to interact with USP11, 15, and 19 and these were found to cooperate in the deubiquitination of poly-ubiquitinated TβRI [23].

USP19 is unique among the DUB family with over 100 members in that it contains a carboxy-terminal transmembrane (TM) domain that targets it to the endoplasmic reticulum (ER) with the active site facing the cytosol [26]. Except for this USP19-ER isoform, USP19 is also expressed as another major isoform without the TM domain, which localises in cytoplasm (herein referred to as USP19-CY) [27]. The USP19-ER and USP19-CY isoforms arise from alternative splicing of the last exon of the *USP19* gene [28]. Structurally, both isoforms possess two CHORD-SGT1 (CS)/P23 domains in their N-termini that serve as cochaperones to Hsp90 [29] and a central USP domain that has the deubiquitinating activity bearing the essential cysteine (C), aspartic acid (D), histidine (H) amino acid residues and a putative ubiquitin-like (UBL) domain as well as a MYND Zn finger domain [30-32]. Multiple studies of USP19 are focused on the USP19-ER variant and its role in the unfolded protein response to rescue ER-associated degradation (ERAD) of substrates [27]. Besides, USP19 deubiquitinates and thereby regulates the stabilities of the ubiquitin ligase KPC1, inhibitors of apoptosis c-IAP1 and c-IAP2, and hypoxia inducible factor 1α (HIF-1α) during hypoxia [33-35]. However, whether the differential location of USP19 impacts TGF-β signaling, and its role in TGF-β-induced EMT, cell migration and invasion in cancer has remained unclear. Therefore, in this study we investigated the roles of two USP19 splicing variants and demonstrated the opposing actions of them in TGF-β-induced responses. We also elucidate the underlying mechanism by which the USP19-CY isoform promotes TGF-β signaling via deubiquitinating and stabilizing the TβRI, thereby enhancing EMT, cell migration and invasion. However, USP19-TM inhibits TGF-β-induced responses in a DUB activity independent manner by restraining TβRI in the ER.

Materials/Subjects and Methods

Cell culture

HEK 293T cells, human lung adenocarcinoma A549-VIM-RFP cells, MDA-MB-231 breast cancer cells and human osteosarcoma U2OS cells were originally obtained from American Type Culture Collection (ATCC) and cultured in Dulbecco's modified Eagle medium (DMEM, 11965092, Thermo) with 10% fetal bovine serum (FBS, S1860-500, BioWest) and 100 U/mL penicillin-streptomycin (15140148, Thermo). The breast epithelial cell line MCF10A-Ras was derived from MCF10A cells transformed with Ha-Ras (kindly provided by Dr. Fred Miller (Barbara Ann Karmanos Cancer Institute, Detroit, MI), were cultured in DMEM/F12 (11039047, Thermo) containing L-glutamine with 5% horse serum (26050088, Thermo), 20 ng/mL epidermal growth factor (EGF, 01-107, Merck Millipore), 10 mg/mL insulin (91077C, Sigma), 100 ng/mL cholera enterotoxin (C8052, Sigma), 0.5 mg/mL hydrocortisone (H0135, Sigma), and 100 U/mL penicillin-streptomycin. All cell lines were tested for absence of mycoplasma contamination and were authenticated by short tandem repeat (STR) profiling.

Reagents and antibodies

The splicing modulators used were SRPIN340 (5042930001, Sigma), TG003 (T5575, Sigma), indisulam (SML1225, Sigma), GSK3326595 (GSK, HY-101563, MedChemExpress), T025 (HY-112296, MedChemExpress), URMC-099 (HY-12599, MedChemExpress), herboxidiene (10-1614, Focus Biomolecules) and Sudemycin D6 (provided by Dr. A.G. Jochemsen, LUMC). Cycloheximide (CHX) was obtained from Sigma (66-81-9). TGF-β3 was generously provided by Dr. A. Hinck (University of Pittsburg, PA). The antibodies used for immunoprecipitation (IP), immunoblotting (IB), and immunofluorescence (IF) are as follows: phosphor-SMAD2 1:1000 (IB: 3108, cell signaling), total-SMAD2 1:1000 (IB: 3103S, cell signaling), USP19 1:1000 (IB, IF: ab189518, Abcam), GAPDH 1:1000 (IB: MAB374, Millipore), Tubulin 1:1000 (IB: 2148, cell signaling), E-cadherin 1:1000 (IB: 610181, BD Biosciences), N-cadherin 1:1000 (IB: 610920, BD Biosciences), vimentin 1:1000 (IB: 5741, cell signaling), SNAIL 1:1000 (IB: 3879, cell signaling),

vinculin (IB: V9131, Sigma), c-MYC 1:200 (IP: sc-40, Santa Cruz), FLAG 1:1000 (IB: F3165), HA 1:1000 (IB: 1583816, Roche), TβRI 1:1000 (IB: sc-398, Santa Cruz), calnexin 1:1000 (IF: ab22595, Abcam), Alexa Fluor 555 secondary antibody 1:250 or 1:1000 (IF: A-31572, Thermo), Alexa Fluor 488 secondary antibody 1:1000 (IF: A-11001, Thermo).

The antibodies of USP19-CY and USP19-ER were raised in rabbits and purified by Eurogentec. The following USP19-CY and USP19-ER-derived peptide sequences (coupled to Keyhole Limpet Hemocyanin (KLH) antigen carrier) were used for immunization: H-CPEVAPTRTAPERFAP-NH$_2$ and Ac-WVGPLPRGPTTPDEGC-NH$_2$, respectively. Per peptide two rabbits were used, in 28 days a total of three injections were done. Pre-immune, medium bleed and large bleed sera were collected. The enzyme-linked immunosorbent assay (ELISA) was performed by the company to analyze the quality of the antibodies, and the results are shown in Figure S7.

Cloning, transfection, lentiviral infection and generation of stable cell lines

The primers and plasmids used for cloning are listed in Supplementary Table S1. Constructs containing the human USP19-ER and the enzymatic inactive mutant USP19-ER-C506S (CS) were a gift from Yihong Ye (Addgene plasmids 78597 and 78584) [26]. The plasmid containing human USP19-CY was made using the USP19-ER plasmid and a MYC-USP19 plasmid (without the TM domain) obtained from Novartis. The active site mutant USP19-CY-C506A (CA) was generated by site-directed mutagenesis. All these cDNAs were inserted into the pLV-CMV-IRES-PURO lentiviral vector. The human HA-TβRI-KDEL plasmid was made using pcDNA3-HA-TβRI (Addgene plasmid 80876) [36] and BFP-KDEL (a gift from Gia Voeltz, Addgene plasmid 49150) [37].

The lentivirus was produced as previously described [23]. The USP19-CY and USP19-ER lentiviral short hairpin (sh)RNAs were obtained from Sigma (MISSION shRNA library) and the most effective shRNAs, sh-USP19-CY (TRCN0000051713, 5'-GCGTGATTTGATTCTGTTGTA-3') and sh-USP19-ER (TRCN0000371018, 5'-GGCCATGCCTGCCT-

TTGTTGT-3') were used. To generate stable cell lines, cells were infected with a 1:1 dilution of the lentivirus in DMEM complemented with 5 ng/mL of Polybrene (Sigma) and selected with puromycin for one week, and subsequently cultured in the presence of puromycin to maintain selection pressure.

ELISA

A coating solution with USP19-CY or USP19-ER peptide plus control carrier keyhole limpet hemocyanin (KLH) was added to a 96 wells plate at a concentration of 100 ng/well and was incubate for 16 hours (h) at 4 °C, after this the plate was blocked using 1 mg/mL bovine serum albumin (BSA) for 2 h at room temperature. Next, the pre-immune serum and large bleed with various dilution (100 x-218700 x) were added into designated wells and were incubated for 2 h at room temperature. Anti-rabbit-IgG-horse radish peroxidase (HRP) detection antibody was diluted at 1:2500 in phosphate buffered saline (PBS) and added to the wells for 2 h at room temperature. After adding 0.4 mg/mL of o-phenylenediamine (OPD) for 20 minutes (min) at room temperature, t4M H_2SO_4 was added to stop the reaction. The absorbance was measured at 492 nm within 30 min of adding the stop solution.

Quantitative real-time-polymerase chain reaction (qRT-PCR)

Quantitative real-time-polymerase chain reaction (qRT-PCR) was performed as previously described [38]. The primer sequences for the detection of the specific genes are listed in Table S1. All target gene expression levels were normalised to glyceraldehyde 3-phosphate dehydrogenase (GAPDH). Results are shown as technical triplicates and representative of three independent biological experiments.

Ubiquitination, immunoprecipitation, and immunoblotting

HEK293T cells were transfected with Myc-tagged constitutively active TβRI (Myc-caTβRI), HA-ubiquitin (HA-Ub) and indicated constructs for 48 h and treated with 5 µM proteasome inhibitor MG132 (474787, Sigma) for 6 h. Next the cells were lysed in 1% sodium dodecyl sulphate (SDS)-RIPA buffer (25 mM Tris-HCl, pH7.4, 150 mM NaCl, 1% NP40, 0.5% sodium deoxycholate, and 1% SDS) supplemented with protease inhibitors (11836153001, Roche) and 10 mM NEM for 10 min on ice. The

lysates were centrifuged at 11×10^3g for 10 min at 4 °C, and the protein concentrations were then measured using the DC protein assay (Pierce). Thereafter, the lysates were boiled for 5 min to eliminate the detection of ubiquitination of co-immunoprecipitating proteins and diluted to 0.1% SDS in a RIPA buffer. The lysate was then incubated with an anti-Myc antibody overnight after which protein G-Sepharose (GE Healthcare Bio-Sciences AB) was added for 2 h at 4 °C (rotating). After washing beads with (SDS)-RIPA buffer, sample buffer was added to the beads and followed up with immunoblotting analysis. For the immunoprecipitation assay, equal amounts of protein were incubated with anti-Flag agarose beads for 2 hours at 4 °C (rotating). Thereafter, beads were washed five times with TNE buffer at 4 °C and after adding sample buffer were boiled for 5 min. The immunoprecipitated proteins were then separated by SDS polyacrylamide gel electrophoresis (PAGE). Immunoblotting was performed as previously described [38]. All experiments were performed with biological triplicates, and representative results are shown.

TAMRA (carboxytetramethylrhodamine)-ubiquitin-vinyl methyl ester (VME) probe assay

The TAMRA-VME probe assay was carried out as described before [39]. Briefly, HEK293T cells transfected with USP19-CY-wt, USP19-CY-CA, USP19-ER-wt or USP19-ER-CS were lysed in TAMRA ABP buffer (50 mM Tris-HCl, pH 7.4, 250 mM sucrose, 5 mM MgCl$_2$, 1 mM DTT, 0.5% zwitterionic surfactant CHAPS and 0.1% nonyl phenoxypolyethoxyl-ethanol (NP40) supplemented with protease inhibitors. Then samples were sonicated with 5 cycles of 30 seconds on and 30 seconds off on ice. Thereafter, cell lysates were centrifuged at 16×10^3 g for 15 min at 4 °C and the supernatant was transferred to a fresh Eppendorf tube to determine the protein concentration. 1 µM carboxytetrame-thylrhodamine ubiquitin-vinyl methyl ester (TAMRA-Ub-VME) probe (UbiQ-050; UbiQ) was used to label the 25 µg protein extract with the total volume of 25 µL for 30 min at room temperature. Labeling reactions were terminated with sample buffer and heating to 100 °C for 10 min. Labeled proteins were separated by NuPAGE 4-12% Bis-Tris protein gels (WG1402BOX; Invitrogen) and the fluorescence signals were detected using the Typhoon

FLA 9500 Molecular Imager (GE Healthcare) at 550 nm excitation and 590 nm emission.

Transcriptional response assay

The SMAD3/4-dependent $CAGA_{12}$-transcriptional luciferase reporter assay was performed as described before [38]. Briefly, HEK293T cells were transfected with $CAGA_{12}$-luc reporter, β-galactosidase encoding plasmids and indicated plasmid using PEI for 24 h. Then the cells were serum starved for 8 hours and treated with or without TGF-β (1 ng/ml) overnight. The $CAGA_{12}$-mediated transcriptional activity was normalized to β-galactosidase expression. All the experiments were performed with biological triplicates, and representative results are shown.

Dynamic detection of RFP-vimentin expression assay

A549-VIM-RFP cells (in which the red fluorescent protein coding region is cloned in frame in the endogenous Vimentin gene locus [40]) were used to determine the EMT process by detecting the dynamic change in red fluorescent protein (RFP)-tagged vimentin expression. Cells with indicated plasmids (pLKO-EV, sh-CY or sh-ER) were cultured in a 96-well plate in the IncuCyte live cell imaging system and treated with vehicle control or TGF-β (2.5 ng/ml) for the indicated time points. The RFP signals were captured every 4 h over a period of 58 h using a 10 × objective. Then the RFP-vimentin intensity was analyzed by the IncuCyte software and normalised by the RFP signals at 0 h for each group. All the experiments were performed with biological triplicates, and representative results are shown.

IncuCyte and transwell migration assay

MDA-MB-231 cells with indicated plasmids (pLKO-EV, sh-CY or sh-ER) were seeded in the IncuCyte 96-well Essen ImageLock plate (4379, Essen BioScience) and scratched using the IncuCyte WoundMaker (Essen BioScience). The scratched cells were washed with PBS and then cultured in the IncuCyte live cell imaging system. Images were acquired every 2 h over a 14-20 h period using a 10× objective. Relative wound density was analyzed by the IncuCyte cell migration software for each well.

Transswell assays were performed in 24-well invasion chambers with 8.0 µm polyethylene terephthalate membrane (354483, Corning). MDA-MB-231 cells with overexpressed USP19-ER-wt or USP19-ER-CS were serum starved overnight and then seeded into the Transwell inserts with DMEM containing 10% FBS in the lower part of the chamber. Cells inside the chamber were carefully removed by a cotton tip humidified with PBS and the migrated cells were fixed in 4% paraformaldehyde (PFA, 28908, Thermo Fisher Scientific) for 10 min. These migrated cells were stained with 0.5% crystal violet for 30 min. Five random fields were selected and photographed for each condition and the number of cells were counted using ImageJ. All experiments were performed with biological triplicates, and representative results are shown.

Nano-Glo HiBiT lytic detection assay

MDA-MB-231 cells stably expressed green fluorescent protein (GFP) were generated as previously described [41]. The HiBiT tag, a small 11 amino acid peptide [42], was knocked in using CRISPR/Cas9 at the endogenous locus of TβRI, resulting a TβRI in which the HiBiT sequence is inserted at the carboxy terminal of the signal peptide. This cell line allows for the specific detection and quantification of the TβRI at cell surface expression by addition of large BiT (LgBiT) to the cell medium. The HiBiT-TβRI cell line was infected with pLV-EV, USP19-ER-wt or USP19-ER-CS lentivirus and then seeded into a 384 wells plate (781098, Greiner Bio-one). After adhering overnight, the medium was removed and replaced with the PBS/LgBiT/NanoGlo substrate mixture using the NanoGlo-HiBiT Detection kit (N2420, Promega). Cells were incubated with the substrate mixture for 15 min and the signals were measured using the VICTOR multilabel plate reader (2030-0050, PerkinElmer). Thereafter, the plate was imaged in the IncuCyte live cell imaging system to determine the GFP intensity, which is a proxy for the number of live cells. NanoGlo signals were normalized to the GFP intensity.

Zebrafish extravasation assay

Zebrafish extravasation assays were performed as previously describe [43]. The experiments were carried out according to the standard guidelines approved by the local Institutional Committee for Animal

Welfare of the Leiden University. The fish were fixed with 4% paraformaldehyde (PFA) four days after injection with mCherry-labeled MDA-MB-231 cells into the Duct of Cuvier and imaged by inverted SP5 confocal microscopy (Leica Microsystems). The cancer cells that invaded into the avascular tail fin area that is rich in collagen were counted (Figure S4C). The experiments were repeated twice in biologically independent experiments, and at least 25 injected embryos were included for quantification.

Formalin-fixed cell line plugs preparation and immunofluorescence staining

To prepare the formalin-fixed cell line plugs for incorporation into paraffin blocks, we used ultra-low gelling temperature (ULGT) agarose (Agarose type IX-A, Sigma) for resuspension medium and a standard agarose (Agarose type I-A, Sigma) as a re-embedding medium. Firstly, HEK293T cells without or with USP19-CY-wt overexpression were fixed with 10% formalin for 3 h at 4 °C and thereafter centrifuged for 30 seconds at 12×10^3 g. The supernatant was discarded. Then cells were resuspended with 50 µL of 3% ULGT agarose solution and centrifuged for 30 seconds at 12×10^3 g. After removal of the supernatant, the compact agarose cell button was solidified for 10 min at 4 °C. Thereafter, the cell button was removed into the cap of an Eppendorf tube, which was further filled with the standard agarose solution. Following the solidification of the standard agarose gel at room temperature for 2 min, the agarose cell button was placed in a tissue cassette and then subjected to routine tissue processing using an automated tissue processor machine and embedded in paraffin. Then this cell line plug was sectioned and mounted on a slide for immunofluorescence staining, which was performed as same protocol as the IF staining with patient tissues.

The formalin-fixed paraffin-embedded microarrays of breast cancer tissues were purchased and included matched breast cancer and cancer adjacent tissues (BR804b, Biomax), and breast cancer tissues in different stages (IIA, IIB, IIIA and IIIB, BC081116e, Biomax). Both tissue arrays were used for immunofluorescence staining. The tissue microarrays were baked overnight at 37 °C and then for 2 h at 60 °C until the paraffin melted. The slides were then placed in a histoclear bath for 7 min three times.

Thereafter, the slides were rehydrated in fresh absolute ethanol for 7 min twice and transferred once through 90%, 70%, and 50% ethanol, respectively, for 3 min each and washed with Milli Q water for 7 min twice. The slides were boiled in an antigen unmasking buffer (1.5M Tris, pH 8.0, 0.5M EDTA, 10% Tween-20) for 35 min using a pressure cooker. The tissue microarrays were then washed twice with Milli Q water for 5 min and once with PBS. Thereafter, the slides were blocked using 1% BSA diluted in PBS/0.1% Tween for 30 min and incubated with a 1:100 diluted primary USP19-CY antibody in PBS/Tween containing 1% BSA overnight at 4 °C. Thereafter, a 1:250 diluted Alexa Fluor 555 secondary antibody in PBS/Tween/BSA was added onto the tissue arrays and incubated for 2 h at room temperature. Subsequently, the slides were washed twice with PBS/Tween. The slides were then incubated with DAPI (diluted 1:1000 in PBS) for 10 min and then washed twice with PBS/Tween. Prolong Gold antifade Mountant (P36930, Thermo) was used to mount the slides. The stained tissue arrays were imaged using a ZEISS Axio Scan Z1 slide scanner. The percentage of USP19-CY in each of the tissues in the arrays was analyzed using QuPath software. The analysis of the tissue sections was performed in an unbiased blinded manner.

The IF staining of cell lines were performed as described before [38]. The experiments were performed with biological triplicates, and representative results are shown.

Statistical analysis

Statistical analyzes were performed using the Student's unpaired t-test using Prism 8 software (GraphPad La Jolla, CA) or as indicated in the legends. All the data were expressed as mean $\pm$ SD with three biological replicates. P-value is indicated by asterisks in the figures: $*P \leqslant 0.05$, $**P < 0.01$, $***P < 0.001$, $****P < 0.0001$. $P \leqslant 0.05$ was considered statistically significant.

Results

USP19-CY promotes TGF-β signaling, while the USP19-ER shows the opposite effect

USP19 is expressed in cells as two major distinct isoforms arising from the alternative splicing of the 3' terminal exon [27]. The USP19-ER isoform contains the TMD that targets USP19 to the ER membrane with the active site facing the cytosol. This TMD is not present in USP19-CY isoform that is localized in the cytoplasm (Figure 1A). Using immunofluorescent staining in U2OS cells, we confirmed previous observations that USP19-ER is an ER-anchored protein that colocalized with an ER protein calnexin, conversely, USP19-CY showed a cytoplasmatic and plasma membrane localization without the colocalization with calnexin using immunofluorescent staining in U2OS cells (Figure 1B, S1C). For specific detection and depletion of USP19 splice variants, we designed primers, shRNAs and antibodies based on the different cDNA and encoded C-terminal sequences of USP19-ER and USP19-CY (Figure S1A, S1B). USP19 is a member of ubiquitin-specific protease family, we confirmed that USP19-ER and CY variants have deubiquitinating activity using the TAMRA-VME probe assay. Both ER-wt, CY-wt, but not the inactive ER-CS and CY-CA mutants are capable of covalently interacting with the TAMRA-VME probe (Figure S1D). Then we investigated the role of USP19-CY in regulating TGF-β signaling. Consistent with our previous report [23], we found that ectopic expression of USP19-CY-wt promoted the TGF-β-induced SMAD3/4-dependent transcriptional response (Figure 1C). Interestingly, overexpression of USP19-CY-CA mutant significantly inhibited this TGF-β-induced response (Figure 1C). To further validate this result, we generated MDA-MB-231 stably expressing FLAG-tagged USP19-CY-wt or USP19-CY-CA. The ectopic expression of USP19-CY was confirmed at mRNA and protein level (Figure S2A and 1D). MDA-MB-231 cells expressing USP19-CY-wt displayed significantly enhanced TGF-β-induced SMAD2 phosphorylation, conversely, ectopic expression of USP19-CY-CA mutant failed to upregulate TGF-β-induced p-SMAD2 levels (Figure 1D, Figure S2B). Moreover, the USP19-CY-wt-induced upregulation of p-SMAD2 was also observed in HEK293T cells transfected with the control

plasmid (pRK5), CY-wt or CY-CA in the presence of TGF-β (Figure S2C). In line with this notion, ectopic expression of CY-wt significantly enhanced the transcript level of TGF-β target genes including *CCN2* (encoding the protein CTGF) and *SERPINE1* (encoding the protein PAI1) with TGF-β treatment for 6 h, but not the CY-CA mutant in MDA-MB-231 cells (Figure 1E). Thus, USP19-CY promotes TGF-β/SMAD signaling in a DUB-dependent manner. Moreover, MDA-MB-231 cells in
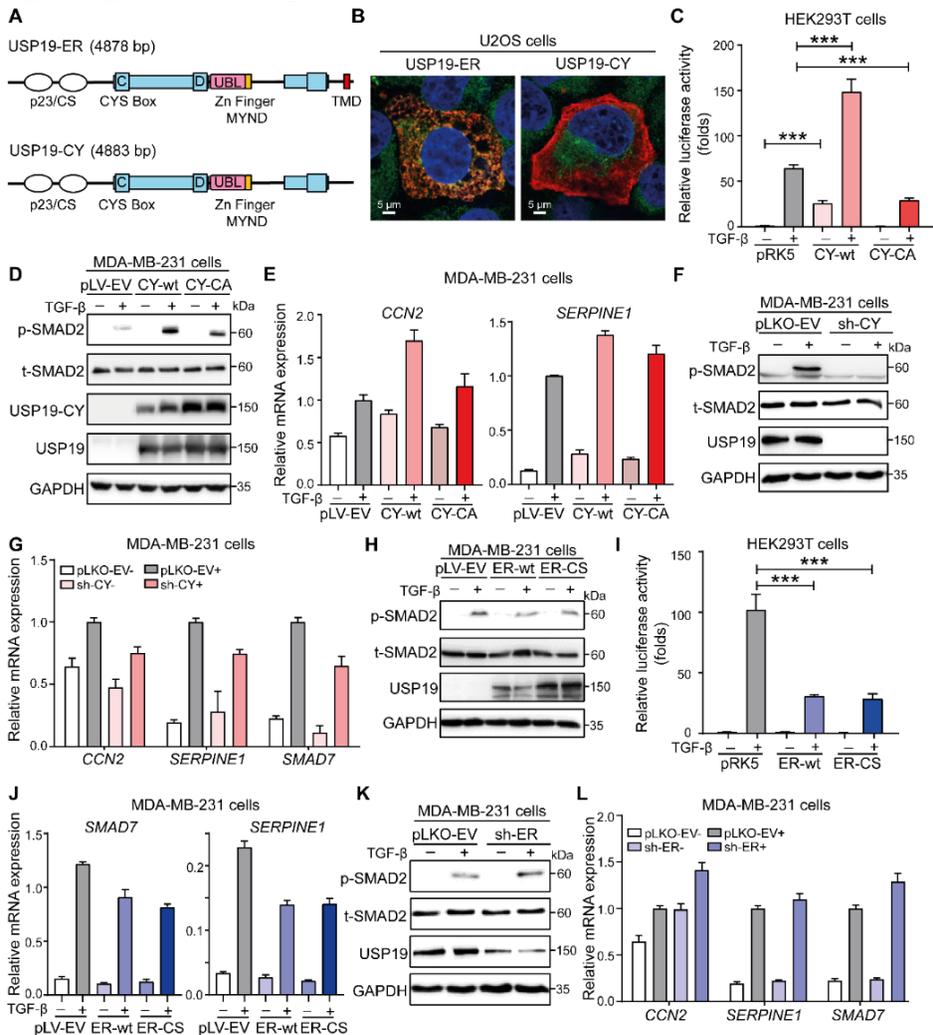

Figure 1. USP19 cytosolic isoform (USP19-CY) promotes TGF-β signaling, conversely, the endoplasmic reticulum (ER)-localized USP19 isoform (USP19-ER) inhibits this TGF-β pathway. (A) Schematic structures depicting USP19-ER and USP19-CY isoforms with common structural domains including a catalytic domain

bearing the essential cysteine (C), aspartic acid (D) and histidine triad of amino acid residues required for catalysis, and unique C-terminal regions. The C-terminal transmembrane domain (TMD) confers ER localisation to the USP19-ER isoform. The catalytic domain also bears a putative ubiquitin-like (UBL) domain as well as a MYND Zn finger domain involved in protein-protein interactions. **(B)** Immunofluorescence analysis of the localization of USP19 (red) and calnexin (green) in U2OS cells transfected with FLAG-tagged wild type USP19-CY and USP19-ER expression plasmids. Nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI, blue). Images were captured with confocal microscopy. Scale bar = 5 μm. **(C)** Effect of USP19-CY-wt or USP19-CY-CA on SMAD3-dependent CAGA$_{12}$-luciferase transcriptional response induced by TGF-β (2.5 ng/mL; overnight treatment) in HEK293T cells. Data is expressed as mean ± SD, n=3. ***, P < 0.001, based on an unpaired Student's t-test. **(D)** Immunoblot analysis of p-SMAD2, total (t)-SMAD2, USP19-CY and total USP19 levels in MDA-MB-231 cells that were infected with empty vector (pLV-EV), wild type USP19-CY (CY-wt) or USP19-CY enzyme inactive mutant (CY-CA) after stimulation of vehicle control or TGF-β (2.5 ng/mL) for 1 h. GAPDH, loading control. **(E)** qRT-PCR analysis of TGF-β target genes, i.e., *CCN2* and *SERPINE1*, in MDA-MB-231 cells stably infected with pLV-EV, CY-wt, CY-CA in the presence of vehicle control or TGF-β (2.5 ng/mL) for 6 h. **(F)** Western blot analysis of p-SMAD2, t-SMAD2 and USP19 levels in MDA-MB-231 cells without or with shRNA-mediated specific knock down of USP19-CY (sh-CY) treated with vehicle control or TGF-β (2.5 ng/mL) for 1 h. GAPDH, loading control. **(G)** Expression levels of TGF-β target genes, i.e. *CCN2*, *SERPINE1* and *SMAD7* in pLKO-EV control or USP19-CY deficient MDA-MB-231 cells treated with vehicle control or TGF-β (2.5 ng/mL) for 6 h. **(H)** Immunoblot analysis of p-SMAD2, t-SMAD2 and USP19 in MDA-MB-231 cells infected with pLV-EV, wild type USP19-ER (ER-wt) and USP19-ER enzyme inactive mutant (ER-CS) and treated with vehicle control or TGF-β (2.5 ng/mL) for 1 h. GAPDH, loading control. **(I)** Measurement of the SMAD3-dependent CAGA$_{12}$-luciferase transcriptional activity induced by TGF-β (2.5 ng/mL) overnight in HEK293T cells that were transfected with ER-wt or ER-CS or pLV-EV expression plasmids. Data was expressed as mean ± SD, n=3. ***, P < 0.001, based on an unpaired Student's t-test. **(J)** qRT-PCR analysis of TGF-β target genes, i.e. *SERPINE1* and *SMAD7*, in MDA-MB-231 cells stably infected with pLV-EV, ER-wt or ER-CS in the presence of vehicle control or TGF-β (2.5 ng/mL) for 6 h. **(K)** Immunoblot of p-SMAD2, t-SMAD2 and USP19 levels in MDA-MB-231 cells without or with shRNA-mediated knock down of USP19-ER (sh-ER) treated with vehicle control or TGF-β (2.5 ng/mL) for 1 h. GAPDH, loading control. **(L)** Expression levels of TGF-β target genes, i.e., *CCN2*, *SERPINE1* and *SMAD7* in PLKV-EV control or USP19-ER deficient MDA-MB-231 cells treated with vehicle control or TGF-β (2.5 ng/mL) for 6 h.

which USP19-CY mRNA and protein was specifically depleted (Figure 1F, Figure S2D), strongly inhibited the TGF-β-induced p-SMAD2 levels (Figure 1F, Figure S2E). This inhibition of SMAD2 phosphorylation was also found in other cell lines with USP19-CY depletion including the

MCF10A-Ras cells and A549-VIM-RFP cells (Figure S2F, S2G, S2H, S2I). Besides, upon shRNA-mediated USP19-CY depletion in MDA-MB-231 cells, MCF10A-Ras cells and A549-VIM-RFP cells, the TGF-β-mediated induction of expression of target genes, including *CCN2*, *SERPINE1* and *SMAD7* were decreased (Figure 1G, Figure S2J, S2K).

Next, similar assays were performed to investigate the role of USP19-ER on TGF-β signaling. After validating the ectopic expression efficiency of ER-wt and ER-CS mutants in MDA-MB-231 cells by qRT-PCR analysis (Figure S3A), we found a significantly decrease of p-SMAD2 levels in cells with ER-wt or ER-CS in the presence of TGF-β (Figure 1H, S3B). This inhibition on TGF-β signaling was also detected in HEK293T cells transfected with ER-wt or ER-CS (Figure S3C). Furthermore, the overexpression of ER-wt or ER-CS strongly suppressed the TGF-β-induced SMAD dependent transcriptional luciferase reporter activity and transcript levels of TGF-β target genes i.e., *SERPINE1* and *SMAD7*. Thus, these results indicate that USP19-ER, in contrast to USP19-CY, inhibits TGF-β signaling and that the catalytic activity of USP19-ER is not needed in this process (Figure 1I, 1J). Consistent with this notion, the specific depletion of USP19-ER in MDA-MB-231 or A549-VIM-RFP cells promoted the TGF-β-induced p-SMAD2 levels (Figure 1K, Figure S3D, S3E and S3F, S3G). In line with this, knockdown of USP19-ER upregulated the expression levels of TGF-β-induced target genes including *CCN2*, *SERPINE1* and *SMAD7* in MDA-MB-231 cells and A549-VIM-RFP cells (Figure 1L, Figure S3H). Taken together, these results indicate that the catalytic activity of USP19-CY is required in promoting TGF-β signaling, while USP19-ER inhibits TGF-β signaling in a DUB activity independent manner.

The USP19-ER isoform inhibits TGF-β-induced EMT and cell migration

To examine the effect of USP19-ER misexpression on EMT, we analyzed the changes of EMT markers in A549-VIM-RFP cells with USP19-ER depletion. The shRNA-mediated knockdown of USP19-ER decreased the expression of the epithelial marker E-cadherin but increased the expression of mesenchymal markers including N-cadherin, vimentin and

SNAIL in the absence or presence of exogenous TGF-β (Figure 2A). Consistent with the immunoblot analysis, the depletion of USP19-ER downregulated the transcript level of epithelial marker *CDH1* (encoding the protein E-cadherin), while promoted the mRNA expression levels of
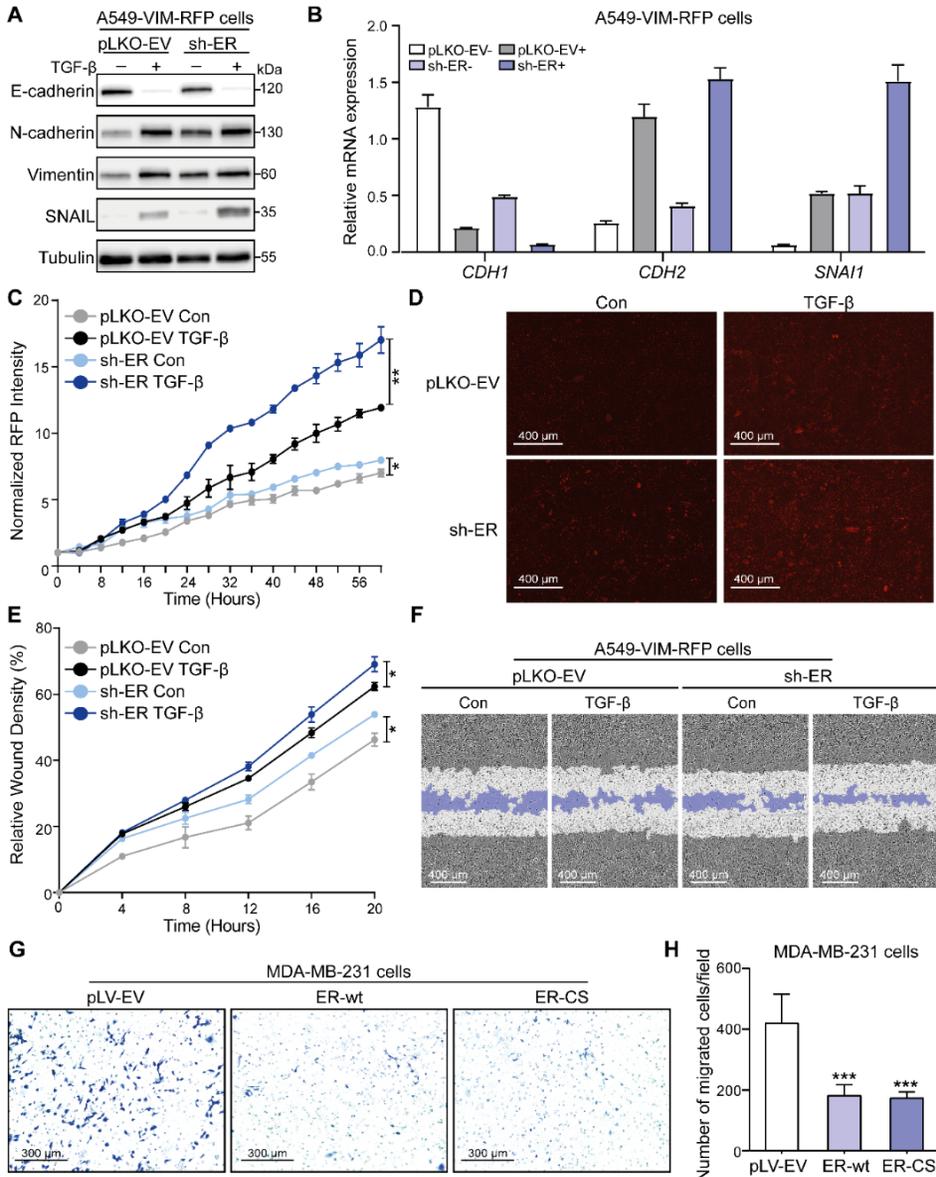

Figure 2. USP19-ER isoform inhibits TGF-β-induced EMT, and cell migration. (A) Immunoblot analysis of epithelial marker E-cadherin, mesenchymal markers N-cadherin, vimentin and SNAIL, and USP19 expression levels in USP19-ER deficient A549-VIM-
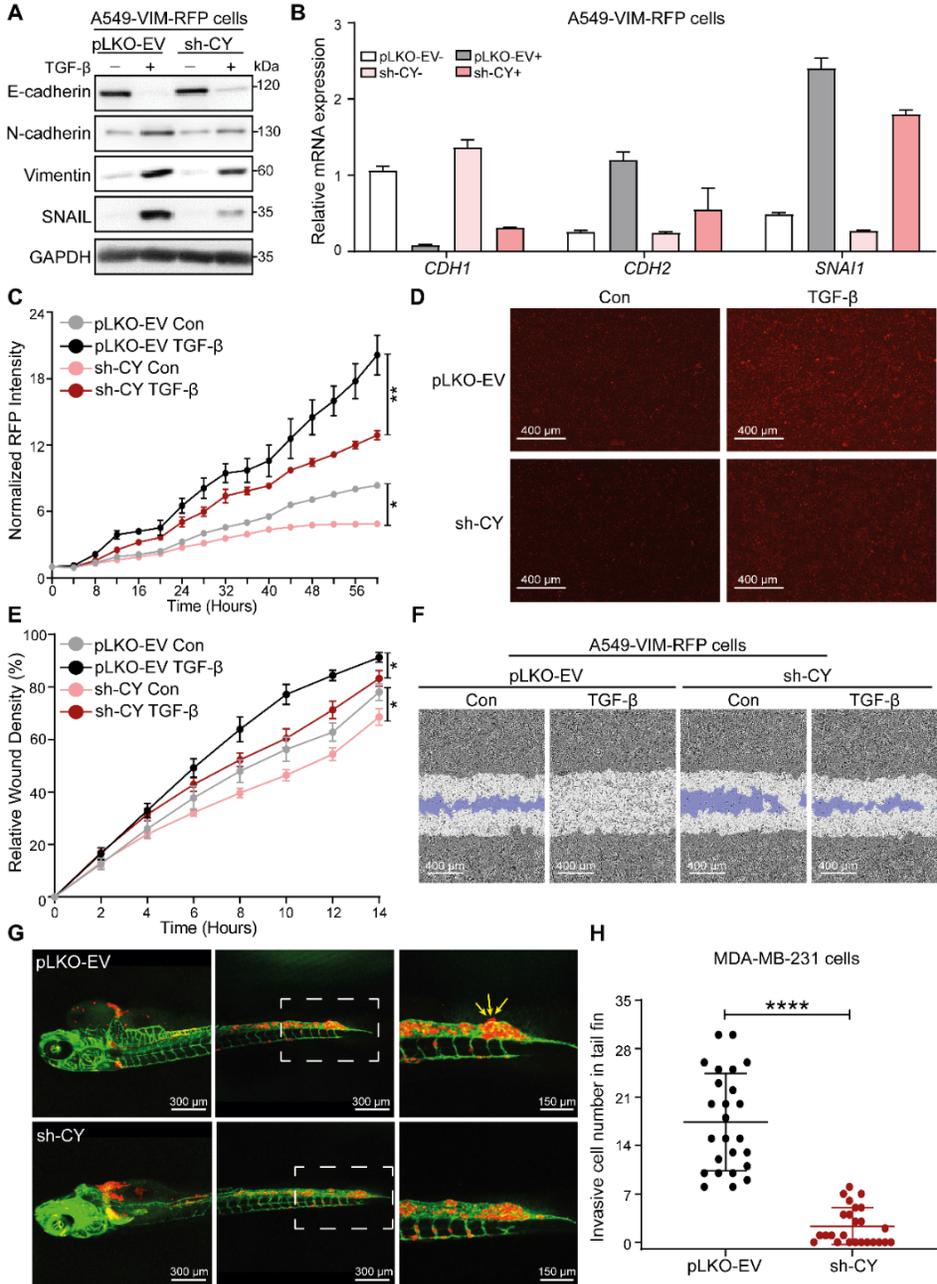
RFP cells that were treated with vehicle control or TGF-β (2.5 ng/mL) for 2 d. Tubulin, loading control. **(B)** qRT-PCR analysis of EMT marker genes *CDH1* (encoding the protein E-cadherin), *CDH2* (encoding the protein N-cadherin) and *SNAI1* (encoding the protein SNAIL) in A549-VIM-RFP cells in pLKO-EV or USP19-ER shRNA infected cells in the presence of TGF-β (2.5 ng/mL) for 2 d. **(C)** Effect of USP19-ER knockdown on vimentin expression in A549-VIM-RFP cells in response to TGF-β (2.5 ng/mL) for the indicated times. Time course for RFP-conjugated vimentin expression was measured by IncuCyte. Red object intensity was normalized by the red intensity at 0 h and expressed as mean ± SD, n=3. **, P < 0.01, based on an unpaired Student's t-test. **(D)** Representative images of RFP-vimentin expression at the end time point (58 h) in A549-VIM-RFP cells with pLKO-EV, sh-ER. Scale bar = 400 μm. **(E)** Real-time scratch assay results as analyzed by IncuCyte in pLKO-EV control or USP19-ER depleted A549-VIM-RFP cells treated with vehicle control or TGF-β (2.5 ng/mL) for the indicated times. Relative wound density (closure) is presented as the mean ± SD, n=3. *, P ≤ 0.05, based on an unpaired Student's t-test. **(F)** Representative images of a scratch wound at the end time point in pLKO-EV control or USP19-ER deficient A549-VIM-RFP cells that were treated with vehicle control or TGF-β (2.5 ng/mL). The region of the original scratch is in white, and the open area of the scratch is indicated in purple. Scale bar = 400 μm. **(G)** Crystal violet staining of cells following transwell migration assay of MDA-MB-231 cells stably infected with pLV-EV, ER-wt, ER-CS plasmids. Scale bar = 300 μm. **(H)** Quantification of the migrated MDA-MB-231 cells stably expressing pLV-EV, ER-wt and ER-CS in the transwell assay. The number of migrated cells per field is shown as mean ± SD, n=5. ***, P < 0.001, based on an unpaired Student's t-test.

mesenchymal markers *CDH2* (encoding the protein N-cadherin) and *SNAI1* (encoding the protein SNAIL) (Figure 2B). The knockdown of USP19-ER-induced promotion of TGF-β-induced EMT was further confirmed by the dynamic increase in RFP-tagged vimentin (Figure 2C, 2D). We next investigated the role of USP19-ER on cell migration. The shRNA-mediated depletion of USP19-ER significantly enhanced both basal and TGF-β-induced A549 cell migration using a scratch assay (Figure 2E, 2F). Similarly, using a transwell assay, we found a lower number of migrated MDA-MB-231 cells that stably overexpressed ER-wt or ER-CS compared to the control group (Figure 2G, 2H). Collectively, these results indicate the negative regulatory role of the USP19-ER isoform (independent of its DUB activity) in TGF-β-induced EMT and cell migration.

The USP19-CY isoform enhances TGF-β-induced EMT, cell migration *in vitro* and invasion *in vivo*

To gain insight into the role of USP19-CY in TGF-β-induced EMT, we first examined the effect of its specific depletion on TGF-β-induced

Figure 3. USP19-CY promotes TGF-β-induced EMT, cell migration and invasion.
(A) Western blot analysis of epithelial marker E-cadherin, mesenchymal markers N-cadherin, vimentin and SNAIL, and USP19 expression levels in A549-VIM-RFP cells without (pLKO-EV) or with USP19-CY knock down that were treated with vehicle control or TGF-β (2.5 ng/mL) for 2 d. Tubulin, loading control. **(B)** qRT-PCR analysis of EMT marker genes *CDH1*, *CDH2* and *SNAI1* in A549-VIM-RFP cells infected with pLKO-EV or sh-CY in absence or presence of TGF-β (2.5 ng/mL) for 2 d. **(C)** Effect of USP19-CY depletion on vimentin expression in A549-VIM-RFP cells treated without or with TGF-β (2.5 ng/mL) for the indicated times. Time course of RFP-conjugated vimentin expression level as measured by IncuCyte. Red object intensity was normalized by the red intensity at 0 h and expressed as mean ± SD, n=3. *, P ≤ 0.05, **, P < 0.01, based on an unpaired Student's t-test. **(D)** Representative images of RFP-vimentin expression at time point (58 h) in A549-VIM-RFP cells with pLKO-EV or sh-CY. Scale bar = 400 μm. **(E)** Real-time scratch assay results as analyzed by IncuCyte in A549-VIM-RFP control cells (pLKO-EV) or with USP19-CY knockdown treated with vehicle control or TGF-β (2.5 ng/mL) for the indicated times. Relative wound density (closure) is presented as the mean ± SD, n=3. *, P ≤ 0.05, based on an unpaired Student's t-test. **(F)** Representative images of scratch wounds at the end time point in pLKO-EV control or USP19-CY depleted A549-VIM-RFP cells that were treated with vehicle control or TGF-β (2.5 ng/mL). The region of the original scratch is in white, and the open area of the scratch is indicated in purple. Scale bar = 400 μm. **(G)** mCherry-labeled MDA-MB-231 cells with pLKO-EV and sh-CY were injected into ducts of Cuvier of zebrafish embryos. Representative images with zoom-in pictures (outlined with a dashed square) of invasive cells were captured 4 days after injection by confocal microscope. Scale bar = 300 μm or 150 μm. Extravasated cells into the avascular collagen-rich tail fin area are indicated with three arrows. **(H)** Quantification of the invasive number in tail fin from 25 embryos for each group. Data are expressed as mean ± SD, n=2. ****, P < 0.0001, based on an unpaired Student's t-test.

EMT marker expression. The shRNA-mediated knockdown of USP19-CY increased the expression of the epithelial marker E-cadherin, but inhibited the expression of mesenchymal markers i.e., N-cadherin, vimentin and SNAIL, both at mRNA and protein levels in A549 cells treated with TGF-β (Figure 3A, 3B). Similarly, shRNA-mediated knockdown of USP19-CY (sh-CY) upregulated E-cadherin and decreased N-cadherin and vimentin were also detected in MCF10A-Ras cells in the presence of TGF-β (Figure S4A). Moreover, we analyzed the dynamic expression changes of RFP-labeled vimentin using the IncuCyte and found that USP19-CY depletion inhibited the vimentin expression both at basal level and in the TGF-β-treated group of A549-VIM-RFP cells (Figure 3C, 3D). In addition, the effect of USP19-CY knockdown on cell

migration was examined by a wound healing assay; USP19-CY depletion significantly downregulated the basal and TGF-β-induced migratory ability of A549 cells (Figure 3E, 3F). To further investigate whether USP19-CY affects the cell invasion, we injected mCherry-labeled MDA-MB-231 cells with pLKO-EV and sh-CY (knockdown efficiency was validated by western blot analysis as shown in Figure S4B) into the ducts of Cuvier (Doc) of a zebrafish embryo (Figure S4C). A significantly lower number of extravascular MDA-MB-231 cells in the tail fin was detected in the USP19-CY depletion group compared to the control group four days after injection (Figure 3G, 3H). These results suggest that the USP19-CY promotes TGF-β-induced EMT, and both basal and TGF-β-mediated cell migration and invasion.

The USP19-ER isoform interacts with and retains TβRI in ER, resulting in the decreased expression of TβRI at the cell membrane

Next, we investigated the mechanism by which USP19-ER inhibits TGF-β/SMAD signaling. As USP19-ER inhibits TGF-β-induced SMAD2 phosphorylation, we hypothesized that USP19-ER may interact with the upstream activator of p-SMAD2, i.e., TβRI. We therefore performed an immunoprecipitation of USP19-ER followed by WB for TβRI in HEK293T cells transfected with FLAG-tagged USP19-ER-wt or USP19-ER-CS and HA-tagged TβRI. We observed that TβRI interacted with both USP19-ER-wt and USP19-ER-CS (Figure 4A). To further validate these results, we analyzed the cell surface expression of endogenous TβRI that was epitope tagged with a HiBiT sequence in MDA-MB-231 cells. We infected these cells with pLV-EV, USP19-ER-wt or USP19-ER-CS lentivirus. The overexpression of USP19-ER (wt and CS) was confirmed by WB using USP19 and USP19-ER antibodies (Figure 4B). The quantification of normalized NanoGlo signals showed lower cell surface TβRI expression in MDA-MB-231 cells with ER-wt and ER-CS as compared to the empty vector group (Figure 4C). Since USP19-ER is a tail-anchored DUB localizing in ER [27], we hypothesized that USP19-ER by interacting with TβRI may restrain TβRI in the ER and thereby interfere with its transportation to the cell membrane. To validate this hypothesis, we generated a construct of TβRI with a C-terminal KDEL sequence for targeting to the ER [44], and performed a CAGA$_{12}$-luciferase

reporter assay in HEK293T cells transfected with empty control (pRK5), wild type TβRI (TβRI-wt), TβRI-KDEL or USP19-ER-wt. Consistent with the expectation, cells transfected with the TβRI-KDEL plasmid showed a significantly decrease of TGF-β/SMAD-induced luciferase activity compared to the control group and cells with the TβRI-wt plasmid, which was comparable to the group of USP19-ER-wt (Figure 4D). Taken together, these findings suggest that USP19-ER inhibits TGF-β signaling in a catalytic activity independent manner by restraining TβRI in the ER (Figure 4E).
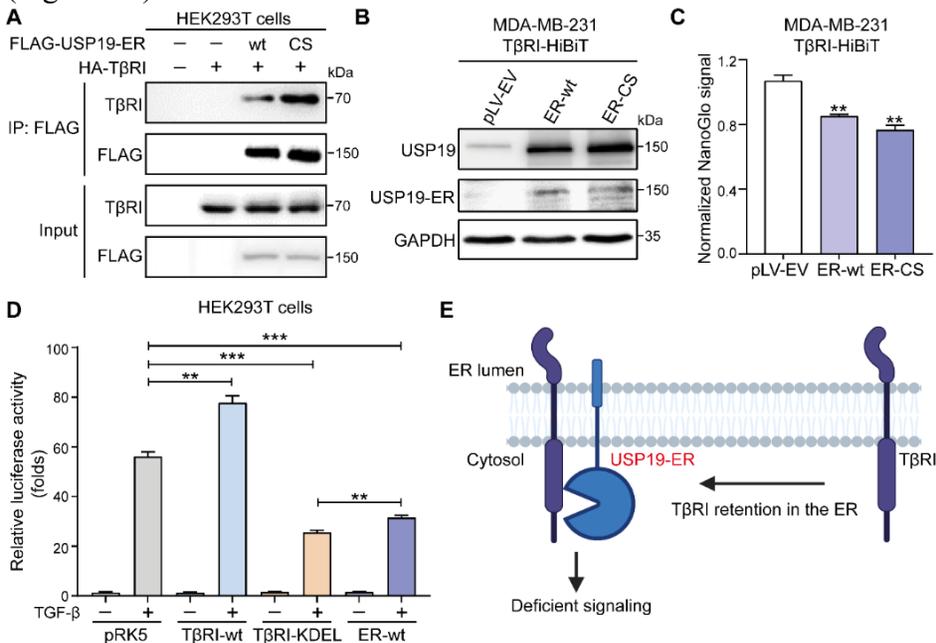

Figure 4. USP19-ER binds to the TβRI and restrains TβRI in ER to decrease its expression on the plasma membrane. (A) The interaction of USP19-ER and TβRI was analyzed by IP of FLAG-tagged USP19-ER (wt or CS mutant) and immunoblot for TβRI in HEK293T cells. **(B)** Western blot analysis of expression levels of USP19 and USP19-ER in MDA-MB-231 cells in which TβRI is endogenously tagged with HiBiT. **(C)** Measurement of TβRI-HiBiT by detecting the NanoGlo signals in MDA-MB-231 cells with pLV-EV, USP19-ER-wt and USP19-ER-CS. Results were normalized by the GFP intensity of cells and expressed as mean ± SD of three independent experiments. **, $P < 0.01$, based on an unpaired Student's t-test. **(D)** Effect of TβRI-wt, TβRI containing the KDEL sequence at carboxy (C)-terminus (TβRI-KDEL) or USP19-ER-wt on $CAGA_{12}$-luciferase transcriptional response induced by TGF-β (2.5 ng/mL) overnight in HEK293T cells. Data was expressed as mean ± SD, n=3. **, $P < 0.01$, ***, $P < 0.001$, based on an unpaired Student's t-test. **(E)** Summary diagram showing the USP19-ER-

mediated inhibition of TGF-β signaling by restraining TβRI in ER and decreasing the amount of TβRI on the cell membrane.

The USP19-CY isoform binds to TβRI, protects TβRI from ubiquitination and increases its stability

To examine the mechanism of USP19-CY-induced promotion of TGF-β/SMAD signaling, we investigated whether USP19-CY targets the TβRI,
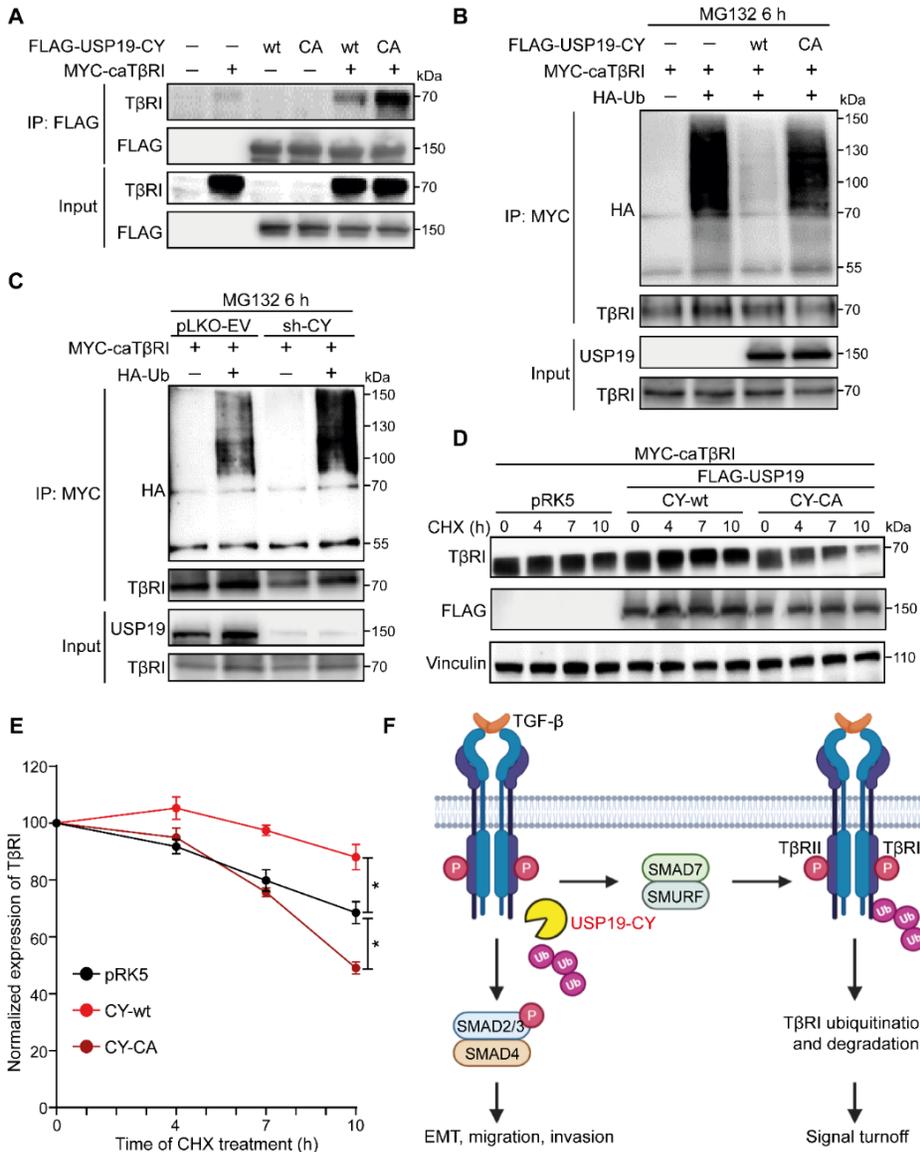
Figure 5. USP19-CY binds to the TβRI and inhibits the ubiquitination and degradation of TβRI. (A) The interaction of USP19-CY and TβRI was analyzed by immunoprecipitation (IP) of FLAG-tagged USP19-CY (wt or CA mutant) and immunoblot for TβRI in HEK293T cells. Ubiquitination of TβRI was detected by IP of MYC-tagged constitutively active TβRI (caTβRI) in HA-Ubiquitin (HA-Ub) transfected HEK293T cells without or with overexpression of CY-wt or CY-CA **(B)** or without (pLKO-EV) or with CY knockdown **(C)**. All groups were treated with MG132 (5 μM) for 6 h. **(D)** Immunoblot analysis of TβRI and FLAG expression levels in HEK293T cells transfected with pRK5, FLAG tagged USP19-CY-wt or USP19-CY-CA expression plasmids in response to cycloheximide (CHX; 50 μg/mL) for the indicated times. Vinculin: loading control. **(E)** Quantification of the TβRI expression levels in pRK5, CY-wt or CY-CA groups of HEK293T cells treated with CHX. Data was normalized to the t=0 controls and expressed as mean ± SD of three independent experiments. *, $P \leqslant 0.05$, based on an unpaired Student's t-test. **(F)** Schematic for USP19-CY-induced promotion of TGF-β signaling by deubiquitinating and increasing the TβRI stability.

as USP19-CY stimulates TGF-β-induced SMAD2 phosphorylation. Both USP19-CY-wt and USP19-CY-CA strongly bound to TβRI when co-expressed in HEK293 cells (Figure 5A). We then examined the effects of USP19-CY-wt and USP19-CY-CA on the ubiquitination of TβRI by overexpressing caTβRI and HA-tagged ubiquitin in HEK293T cells. USP19-CY-wt strongly mitigated the polyubiquitination of caTβRI, conversely, the caTβRI-associated USP19-CY-CA mutant remained highly polyubiquitinated (to similar levels as caTβRI alone) (Figure 5B). Moreover, knockdown of USP19-CY significantly increased the polyubiquitination of TβRI (Figure 5C). The role of USP19-CY in regulating the TβRI stability was studied by examining TβRI expression levels upon ectopic expression in HEK293T cells after treatment with the protein synthesis inhibitor cycloheximide (CHX). The protein half-live of TβRI was prolonged by USP19-CY-wt but not CY-CA (Figure 5D, 5E). These data results suggest that USP19-CY is a DUB for TβRI, which protects TβRI from polyubiquitination and subsequent degradation (Figure 5F).

USP19-CY isoform is highly expressed in breast cancer tissues and while herboxidiene promotes USP19-ER expression, it inhibits USP19-CY expression

We next investigated whether the USP19-CY expression can be linked with the prognosis of breast cancer patients. Therefore, we performed

immunofluorescence staining of the USP19-CY protein level using a USP19-CY-specific antibody. The specificity and efficiency of the antibody was validated by the IF staining of USP19-CY in pRK5 and USP19-CY transfected HEK293T cell line plugs embedded in paraffin (Figure S4D). Then we performed the IF staining of USP19-CY in two tissue microarrays, one contains 34 paired breast cancer tissues and adjacent phenotypically normal tissues derived from 34 patients and another includes the different stages (IIA, IIB, IIIA, IIB) of breast cancer tissues and 10 adjacent normal tissues (Figure 6A, 6B). We observed that USP19-CY levels were higher in breast cancer tissues compared to normal adjacent tissues (Figure 6C, 6D). Furthermore, the more advanced breast cancer tissue stages, i.e. stage IIIA and IIIB demonstrated higher expression of USP19-CY compared to the breast cancer tissue stage IIA and IIB (Figure 6D, Figure S4E).

3

The expression of USP19-ER and USP19-CY isoforms are a result from alternative splicing [27, 45]. Thus, we aimed to identify specific small molecule splicing modulators that favor USP19-ER at expense of USP19-CY expression. We therefore challenged cells with 8 splicing modulators (Table S2) and analyzed USP19-ER versus CY expression. qRT-PCR analysis revealed that T025 and herboxidiene significantly inhibited the expression levels of *USP19-CY* and *USP19*, but increased the *USP19-ER* mRNA levels (Figure S5). Other modulators had no clear effect on transcript levels of *USP19*, *USP19-CY* and *USP19-ER* in HEK293T, A549, MDA-MB-231 and MCF10A-Ras cells (Figure S5). The role of herboxidiene on USP19 splicing was further confirmed at the protein expression levels of the USP19 isoforms. While herboxidiene decreased the USP19 and USP19-CY expression, it upregulated USP19-ER levels in A549 cells and HEK293T cells (Figure 6E, Figure S6A). Challenging cells with T025 revealed the downregulation of USP19 and USP19-CY, but USP19-ER was not significantly changed. Next, we investigated the effect of herboxidiene on TGF-β signaling, EMT and cell migration. We observed a significant inhibition by herboxidiene on the TGF-β-induced p-SMAD2 response in MDA-MB-231 cells with the empty vector (pLV-EV). This inhibition was rescued by ectopic overexpression of USP19-
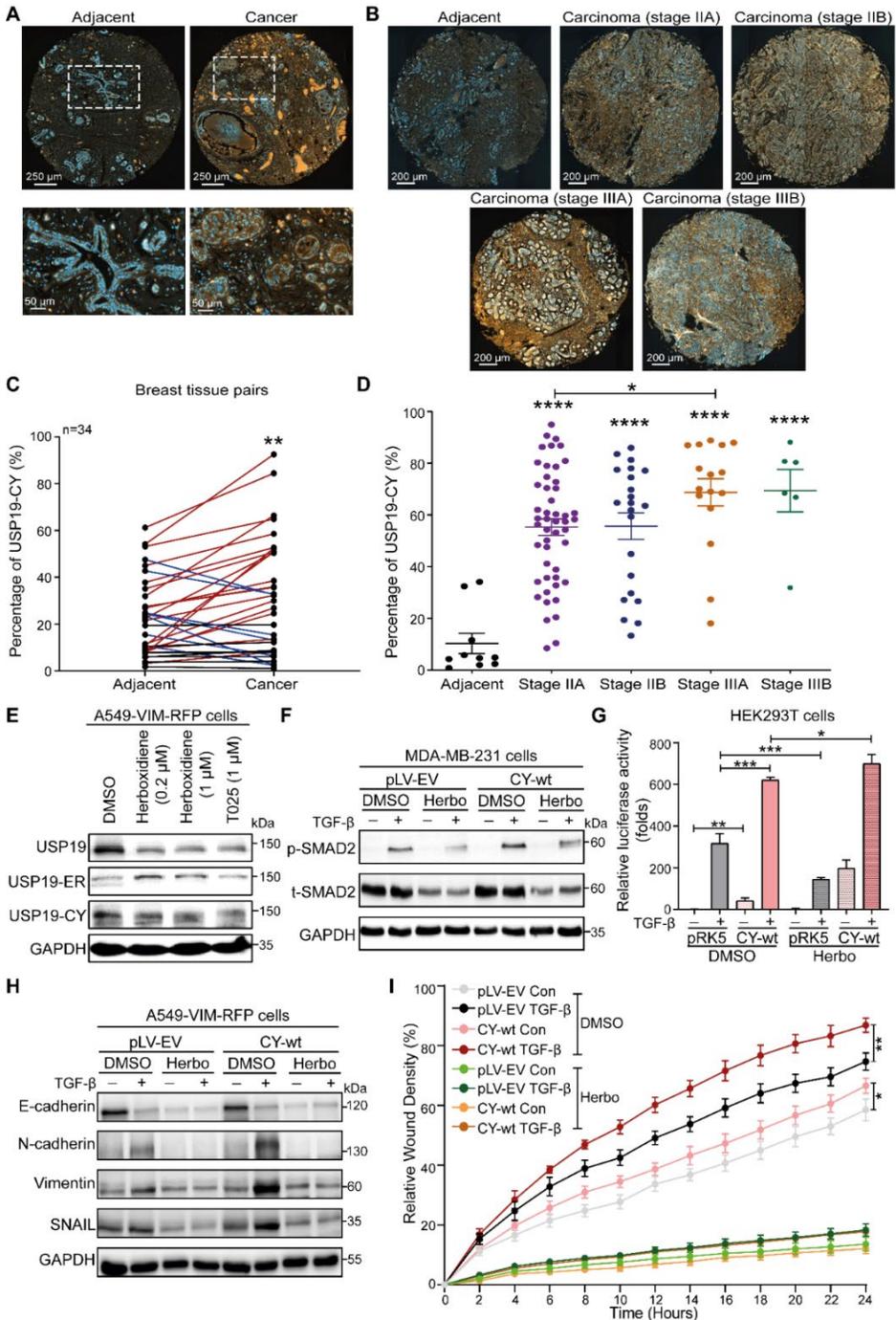
Figure 6. USP19-CY is highly expressed in breast cancer tissues, and USP19 mRNA splicing is regulated by herboxidiene. Representative images of USP19-CY (red)

immunofluorescent staining in human breast cancer tissue microarray containing 34 paired cancer adjacent tissues and cancer tissues **(A)** or cancer tissues of different stages (stage IIA, IIB, IIIA, IIIB) **(B)**. Nuclei were counterstained with DAPI (blue). Large field and zoom-in pictures (outlined with a dotted square) are shown. Scale bar = 250 μm, 50 μm or 250 μm. **(C)** Quantification of the percentage of USP19-CY in breast tissue pairs (adjacent and cancer tissues). Red lines indicate the significant upregulation and blue line indicate the downregulation of USP19-CY in cancer tissues compared to adjacent tissues, black lines indicate no significant change of USP19-CY in tissue pairs. Data are represented as mean ± SD, tissue pairs, n=34, **, $P < 0.01$, based on a paired Student's t-test. **(D)** Quantification of USP19-CY percentages in breast cancer adjacent tissues and cancer tissues of different stages. Data are expressed as mean ± SD, adjacent tissues, n=10; adenocarcinoma (stage IIA), n=49; adenocarcinoma (stage IIB), n=22; adenocarcinoma (stage IIIA), n=16; adenocarcinoma (stage IIIB), n=6; *, $P \leq 0.05$, ****, $P < 0.0001$, based on an unpaired Student's t-test. **(E)** Immunoblot analysis of USP19, USP19-CY and USP19-ER in A549-VIM-RFP cells treated with 0.2 or 1 μM herboxidiene and 1 μM T025. GAPDH, loading control. **(F)** MDA-MB-231 cells stably infected with pLV-EV or USP19-CY-wt were pre-treated with 1 μM herboxidiene (Herbo) for 24 h and then combined with vehicle control or TGF-β (2.5 ng/mL) for 1 h, followed by immunoblot analysis of p-SMAD2 and t-SMAD2 expression levels. GAPDH: loading control. **(G)** HEK293T cells transfected with pRK5 or USP19-CY-wt were pre-treated with 1 μM herboxidiene (Herbo) for 24 h and then combined with vehicle control or TGF-β (2.5 ng/mL) overnight, followed by the analysis of CAGA$_{12}$-luciferase transcriptional responses. Data was expressed as mean ± SD, n=3. *, $P \leq 0.05$, **, $P < 0.01$, ***, $P < 0.001$, based on an unpaired Student's t-test. **(H)** A549-VIM-RFP cells stably infected with pLV-EV or USP19-CY-wt were pre-treated with 1 μM herboxidiene (Herbo) for 24 h and then combined with vehicle control or TGF-β (2.5 ng/mL) for 48 h, followed by immunoblot analysis of epithelial marker E-cadherin and mesenchymal markers N-cadherin, vimentin and SNAIL expression levels. GAPDH: loading control. **(I)** A549-VIM-RFP cells with pLV-EV and USP19-CY-wt plasmids were pre-treated with 1 μM herboxidiene for 24 h and then incubated with vehicle control or TGF-β (2.5 ng/mL) for the indicated times. Time course for scratch assay results as analyzed by IncuCyte. Relative wound density (closure) was presented as the mean ± SD, n=3. *, $P \leq 0.05$, **, $P < 0.01$, based on an unpaired Student' s t-test.

CY-wt (Figure 6F, Figure S6B). Conversely, T025 showed the same inhibition effect of TGF-β-induced SMAD2 phosphorylation in MDA-MB-231 cells stably expressing pLV-EV and UP19-CY-wt (Figure S6C, S6D). In line with this, ectopic expression of USP19-CY-wt significantly rescued the CAGA-luciferase activities in HEK293T cells treated with herboxidiene but not T025 stimulated cells (Figure 6G, Figure S6E). Furthermore, herboxidiene strongly inhibited TGF-β-induced expression of either epithelial marker E-cadherin and mesenchymal markers

including N-cadherin, vimentin and SNAIL, indicating the various mechanism of herboxidiene-regulated EMT in A549 cells (Figure 6H). Furthermore, herboxidiene completely blocked migration of A549 cells also confirmed this notion (Figure 6I). Collectively, our results suggests that USP19-CY is highly expressed in breast cancer tissues. Herboxidiene (not T025) regulates the splicing of USP19 by altering the USP19-CY isoform to the USP19-ER isoform. Consistent with this latter finding, herboxidiene inhibits TGF-β signaling, EMT and cancer cell migration.

Discussion

USP19-CY and USP19-ER share TβRI interaction, but have opposing roles on TGF-β/SMAD signaling

In this study, we observed the opposite roles of two USP19 isoforms in TGF-β signaling and found both of them interact with TβRI. We identified that the USP19-CY isoform promoted TGF-β/SMAD signaling which required its DUB activity. Mechanistically, we showed that USP19-CY directly deubiquitinates and stabilizes TβRI in the plasma membrane. These results are consistent with a previous genetic gain of function screen in which USP19 was identified (among many other cDNAs) to promote the TGF-β-induced SMAD3/4-dependent transcriptional luciferase reporter activity [23]; we confirmed that the USP19 cDNA construct used was the USP19-CY isoform. In contrast to USP19-CY, we showed that the USP19-ER isoform negatively regulated TGF-β/SMAD signaling in a DUB activity independent manner. Furthermore, USP19-ER restrained TβRI in the ER, thereby resulting in lower TβRI levels in the plasma membrane and making the cells less TGF-β responsive. This notion was further validated using TβRI-KDEL fusion construct that targets TβRI to ER. Ectopic expression of TβRI-KDEL displayed a comparable inhibition of TGF-β signaling as USP19-ER. Indeed, "chaperone like" activity of USP19-ER has also been proposed, by which USP19-ER might promote folding by interacting with HSP90 through its CS/p23 domain [46]. This may provide a possible mechanism of how USP19-ER affects the folding of TβRI, resulting in its retention in the ER.

Opposite roles of USP19-ER and USP19-CY on TGF-β-induced EMT, migration and invasion

In the breast and lung cancer cells that we have used in our study, USP19-CY is the major isoform that was always much higher expressed than the USP19-ER isoform. Indeed, overexpression of TGF-β has been demonstrated in human tumor models and is seen clinically in many tumors including cancers of the breast and lung [47, 48]. Thus, the highly expressed USP19-CY isoform and TGF-β may have a potential correlation and affect breast and lung tumorigenesis. In our study, the opposing roles of USP19-ER and USP19-CY on TGF-β/SMAD signaling resulted in that USP19-CY stimulated and USP19-ER inhibited TGF-β-induced biological processes in breast and lung cancer, including TGF-β-induced EMT and cell migration. Importantly, we observed USP19-CY-mediated promotion of extravasation of the breast cancer cells MDA-MB-231 in the zebrafish xenograft model. Of interest, USP19-ER was identified as a negative regulator of proliferation and migration of clear cell renal cell carcinoma (ccRCC) by suppressing ERK map kinase activation [49]. In another study, overexpression of USP19-ER was found to increase breast cancer cell migration and invasion, which is dependent on its catalytic activity [50]. By deubiquitylation of LRP6 by USP19, Wnt signaling was increased and thereby induced cell migration and invasion [50]. In the same study, knockdown of total USP19 inhibited MDA-MB-231 cell migration [50]; this is in line with the depletion of USP19-CY, which is the major isoform in MDA-MB-231 cells, also inhibited cell migration. Furthermore, authors in this study also showed that USP19 depletion decreased tumor growth and metastasis *in vivo*. This is consistent with the critical role of USP19-CY in TGF-β-induced extravasation and metastasis of MDA-MB-231 cell in zebrafish and mice xenograft models.

USP19-CY expression is associated with the poor prognosis of breast cancer patients

Importantly, consistent with a pro-invasive/EMT activity mediated by the USP19-CY variant, we revealed that USP19-CY is higher expressed in breast cancer than phenotypically normal adjacent tissues, and the higher expression level is related to more advanced staged cancer. This suggests

3

that USP19-CY expression can be linked to poor prognosis of breast cancer patients. Our results were confirmed by a previous study in which high expression of USP19 was found to be associated with a significantly lower frequency of distant relapse-free survival in early breast cancer patients [50]. Additionally, elevated USP19 was observed in gastric cancer cells and tissues, and gastric cancer patients with high-level USP19 expression showed poor survival [51]. Although the previous studies didn't specify the USP19 isoforms, these can still offer some clues of positive roles of USP19-CY in tumorigenesis due to its predominantly expression in most cancers. However, an analysis of isoform expression signatures associated with tumor stages in kidney renal clear cell carcinoma (KIRC) showed uc003cvz.3, the cytosolic isoform of USP19, was significantly decreased in patients with stage IV KIRC, whereas higher uc003cvz.3 expression suggested improved survival rates [52]. Therefore, anti- or pro-tumor effects mediated by USP19-CY may differ depending on cancer subtype.

Roles of splicing in cancer progression

Multiple studies have highlighted the frequently alteration of splicing in cancer, and a causal role of splice variant expression contributing to cancer progression [53, 54]. For example, CD44 variant isoforms (CD44v) arising from the inclusion of one or more of the variable exons is expressed in epithelial cells, while the CD44 standard isoform (CD44s) is mainly presented in mesenchymal cells. Thus, pharmacological manipulation of alternative splicing has been pursued for anti-cancer therapeutic benefit. As such, a number of chemical small molecule compounds have been identified to inhibit the core spliceosome or the phosphorylation of splicing regulatory proteins [55]. Notably, we identified herboxidiene as a USP19 splicing modulator by strongly decreasing the expression of USP19-CY but increasing USP19-ER at both protein and mRNA levels in breast and lung cancer cells. Another splicing modulator, T025, had no effect on USP19 isoform ratios, but resulted in the downregulation of the USP19-CY isoform. The herboxidiene (not T025)-induced inhibition of TGF-β signaling can be rescued by the overexpression USP19-CY, which validates the opposite roles of two USP19 isoforms on the signaling. Moreover, we found that herboxidiene

can completely inhibit the basal expression of epithelial marker E-cadherin and TGF-β-induced mesenchymal markers including N-cadherin, vimentin and SNAIL, and migration of lung cancer A549 cells. Inhibition of mesenchymal marker expression may mitigate single cell migration/invasion. The low levels of E-cadherin may have a negative effect on the collective migration of these cancer cells. Besides our observation on splicing of USP19 mRNA, herboxidiene has been reported to regulate the pre-mRNA splicing of p27, a key inhibitor of the cell cycle, leading to the accumulation of spliced p27 and inhibition of cyclin E-Cdk2 complex formation [55]. Taking together, targeting, alternative splicing with compounds such as herboxidiene, has potentials for cancer therapeutics.

In conclusion, our findings have demonstrated the distinct roles of two USP19 isoforms, i.e., USP19-ER and USP19-CY, in regulating TGF-β signaling by targeting TβRI with different mechanisms. USP19-ER-mediated inhibition of TGF-β/SMAD signaling that are causally linked to the decreased TGF-β-induced EMT and migration of breast and lung cancer cells. Oppositely, USP19-CY promotes TGF-β/SMAD-induced breast and lung cancer cell EMT, cell migration and extravasation *in vitro* and *in vivo* models. Moreover, consistent with these findings, USP19-CY expression is linked to poor prognosis of breast cancer patients. The identification of herboxidiene as a specific modulator of USP19 splicing and its concomitant inhibitory effects on TGF-β/SMAD signaling, and cancer migration further validate the opposing roles of USP19-ER and USP19-CY in these responses. It will be interesting to explore the potential of USP19-CY for breast cancer therapy as a prognostic biomarker and its potential as a molecular target either by redirecting its expression towards USP19-ER or inhibiting its deubiquitinating activity with selective small molecules.

Acknowledgements

Competing Interests

The authors declare that they have no conflicts of interest in regard to the contents of this article.

References

1   Colak S, Ten Dijke P. Targeting TGF-β Signaling in Cancer. Trends Cancer 2017; 3: 56-71.

2   Massague J. TGFβ in Cancer. Cell 2008; 134: 215-230.

3   Ikushima H, Miyazono K. TGFβ signalling: a complex web in cancer progression. Nat Rev Cancer 2010; 10: 415-424.

4   Massague J. How cells read TGF-β signals. Nat Rev Mol Cell Biol 2000; 1: 169-178.

5   Shi Y, Massague J. Mechanisms of TGF-β signaling from cell membrane to the nucleus. Cell 2003; 113: 685-700.

6   Heldin CH, Miyazono K, ten Dijke P. TGF-β signalling from cell membrane to nucleus through SMAD proteins. Nature 1997; 390: 465-471.

7   Levy L, Hill CS. Smad4 dependency defines two classes of transforming growth factor β (TGF-β) target genes and distinguishes TGF-β-induced epithelial-mesenchymal transition from its antiproliferative and migratory responses. Mol Cell Biol 2005; 25: 8108-8125.

8   ten Dijke P, Hill CS. New insights into TGF-β-Smad signalling. Trends Biochem Sci 2004; 29: 265-273.

9   Katsuno Y, Lamouille S, Derynck R. TGF-β signaling and epithelial-mesenchymal transition in cancer progression. Curr Opin Oncol 2013; 25: 76-84.

10   Lamouille S, Xu J, Derynck R. Molecular mechanisms of epithelial-mesenchymal transition. Nat Rev Mol Cell Biol 2014; 15: 178-196.

11   Yang J, Antin P, Berx G, Blanpain C, Brabletz T, Bronner M et al. Guidelines and definitions for research on epithelial-mesenchymal transition. Nat Rev Mol Cell Biol 2020; 21: 341-352.

12 Nakajima Y, Yamagishi T, Hokari S, Nakamura H. Mechanisms involved in valvuloseptal endocardial cushion formation in early cardiogenesis: roles of transforming growth factor (TGF)-β and bone morphogenetic protein (BMP). Anat Rec 2000; 258: 119-127.

13 Barriere G, Fici P, Gallerani G, Fabbri F, Rigaud M. Epithelial Mesenchymal Transition: a double-edged sword. Clin Transl Med 2015; 4: 14.

14 Hao Y, Baker D, Ten Dijke P. TGF-β-Mediated Epithelial-Mesenchymal Transition and Cancer Metastasis. Int J Mol Sci 2019; 20.

15 Derynck R, Weinberg RA. EMT and Cancer: More Than Meets the Eye. Dev Cell 2019; 49: 313-316.

16 Lonn P, Moren A, Raja E, Dahl M, Moustakas A. Regulating the stability of TGFβ receptors and Smads. Cell Res 2009; 19: 21-35.

17 De Boeck M, ten Dijke P. Key role for ubiquitin protein modification in TGFβ signal transduction. Ups J Med Sci 2012; 117: 153-165.

18 Komander D. The emerging complexity of protein ubiquitination. Biochem Soc Trans 2009; 37: 937-953.

19 Zhu H, Kavsak P, Abdollah S, Wrana JL, Thomsen GH. A SMAD ubiquitin ligase targets the BMP pathway and affects embryonic pattern formation. Nature 1999; 400: 687-693.

20 Zhang Y, Chang C, Gehling DJ, Hemmati-Brivanlou A, Derynck R. Regulation of Smad degradation and activity by Smurf2, an E3 ubiquitin ligase. Proc Natl Acad Sci U S A 2001; 98: 974-979.

21 Lo RS, Massague J. Ubiquitin-dependent degradation of TGF-β-activated smad2. Nat Cell Biol 1999; 1: 472-478.

22 Liu S, de Boeck M, van Dam H, Ten Dijke P. Regulation of the TGF-β pathway by deubiquitinases in cancer. Int J Biochem Cell Biol 2016; 76: 135-145.

23 Zhang L, Zhou F, Drabsch Y, Gao R, Snaar-Jagalska BE, Mickanin C et al. USP4 is regulated by AKT phosphorylation and directly deubiquitylates TGF-β type I receptor. Nat Cell Biol 2012; 14: 717-726.

24 Iyengar PV, Jaynes P, Rodon L, Lama D, Law KP, Lim YP et al. USP15 regulates SMURF2 kinetics through C-lobe mediated deubiquitination. Sci Rep 2015; 5: 14733.

25 Eichhorn PJ, Rodon L, Gonzalez-Junca A, Dirac A, Gili M, Martinez-Saez E et al. USP15 stabilizes TGF-β receptor I and promotes oncogenesis through the activation of TGF-β signaling in glioblastoma. Nat Med 2012; 18: 429-435.

26 Lee JG, Takahama S, Zhang G, Tomarev SI, Ye Y. Unconventional secretion of misfolded proteins promotes adaptation to proteasome dysfunction in mammalian cells. Nat Cell Biol 2016; 18: 765-776.

3

27   Hassink GC, Zhao B, Sompallae R, Altun M, Gastaldello S, Zinin NV et al. The ER-resident ubiquitin-specific protease 19 participates in the UPR and rescues ERAD substrates. EMBO Rep 2009; 10: 755-761.

28   Wiles B, Miao M, Coyne E, Larose L, Cybulsky AV, Wing SS. USP19 deubiquitinating enzyme inhibits muscle cell differentiation by suppressing unfolded-protein response signaling. Mol Biol Cell 2015; 26: 913-923.

29   Zhang M, Boter M, Li K, Kadota Y, Panaretou B, Prodromou C et al. Structural and functional coupling of Hsp90- and Sgt1-centred multi-protein complexes. EMBO J 2008; 27: 2789-2798.

30   Wing SS. Deubiquitinating enzymes in skeletal muscle atrophy-An essential role for USP19. Int J Biochem Cell Biol 2016; 79: 462-468.

31   Lee JG, Kim W, Gygi S, Ye Y. Characterization of the deubiquitinating activity of USP19 and its role in endoplasmic reticulum-associated degradation. J Biol Chem 2014; 289: 3510-3517.

32   Liu Y, Chen W, Gaudet J, Cheney MD, Roudaia L, Cierpicki T et al. Structural basis for recognition of SMRT/N-CoR by the MYND domain and its contribution to AML1/ETO's activity. Cancer Cell 2007; 11: 483-497.

33   Lu Y, Adegoke OA, Nepveu A, Nakayama KI, Bedard N, Cheng D et al. USP19 deubiquitinating enzyme supports cell proliferation by stabilizing KPC1, a ubiquitin ligase for p27Kip1. Mol Cell Biol 2009; 29: 547-558.

34   Mei Y, Hahn AA, Hu S, Yang X. The USP19 deubiquitinase regulates the stability of c-IAP1 and c-IAP2. J Biol Chem 2011; 286: 35380-35387.

35   Altun M, Zhao B, Velasco K, Liu H, Hassink G, Paschke J et al. Ubiquitin-specific protease 19 (USP19) regulates hypoxia-inducible factor 1alpha (HIF-1alpha) during hypoxia. J Biol Chem 2012; 287: 1962-1969.

36   Moren A, Imamura T, Miyazono K, Heldin CH, Moustakas A. Degradation of the tumor suppressor Smad4 by WW and HECT domain ubiquitin ligases. J Biol Chem 2005; 280: 22115-22123.

37   Friedman JR, Lackner LL, West M, DiBenedetto JR, Nunnari J, Voeltz GK. ER tubules mark sites of mitochondrial division. Science 2011; 334: 358-362.

38   Zhang J, Thorikay M, van der Zon G, van Dinther M, Ten Dijke P. Studying TGF-β Signaling and TGF-β-induced Epithelial-to-mesenchymal Transition in Breast Cancer and Normal Cells. J Vis Exp 2020.

39   Liu S, Gonzalez-Prieto R, Zhang M, Geurink PP, Kooij R, Iyengar PV et al. Deubiquitinase Activity Profiling Identifies UCHL1 as a Candidate Oncoprotein That Promotes TGFβ-Induced Breast Cancer Metastasis. Clin Cancer Res 2020; 26: 1460-1473.

40   Wang W, Douglas D, Zhang J, Kumari S, Enuameh MS, Dai Y et al. Live-cell imaging and analysis reveal cell phenotypic transition dynamics inherently missing in snapshot data. Sci Adv 2020; 6.

41   Deckers M, van Dinther M, Buijs J, Que I, Lowik C, van der Pluijm G et al. The tumor suppressor Smad4 is required for transforming growth factor β-induced epithelial to mesenchymal transition and bone metastasis of breast cancer cells. Cancer Res 2006; 66: 2202-2209.

42   Schwinn MK, Machleidt T, Zimmerman K, Eggers CT, Dixon AS, Hurst R et al. CRISPR-Mediated Tagging of Endogenous Proteins with a Luminescent Peptide. ACS Chem Biol 2018; 13: 467-474.

43   Ren J, Liu S, Cui C, Ten Dijke P. Invasive Behavior of Human Breast Cancer Cells in Embryonic Zebrafish. J Vis Exp 2017.

44   Stornaiuolo M, Lotti LV, Borgese N, Torrisi MR, Mottola G, Martire G et al. KDEL and KKXX retrieval signals appended to the same reporter protein determine different trafficking between endoplasmic reticulum, intermediate compartment, and Golgi complex. Mol Biol Cell 2003; 14: 889-902.

45   He WT, Zheng XM, Zhang YH, Gao YG, Song AX, van der Goot FG et al. Cytoplasmic Ubiquitin-Specific Protease 19 (USP19) Modulates Aggregation of Polyglutamine-Expanded Ataxin-3 and Huntingtin through the HSP90 Chaperone. PLoS One 2016; 11: e0147515.

46   Perrody E, Abrami L, Feldman M, Kunz B, Urbe S, van der Goot FG. Ubiquitin-dependent folding of the Wnt signaling coreceptor LRP6. Elife 2016; 5.

47   Chod J, Zavadova E, Halaska MJ, Strnad P, Fucikova T, Rob L. Preoperative transforming growth factor-β1 (TGF-β1) plasma levels in operable breast cancer patients. Eur J Gynaecol Oncol 2008; 29: 613-616.

48   Li J, Shen C, Wang X, Lai Y, Zhou K, Li P et al. Prognostic value of TGF-β in lung cancer: systematic review and meta-analysis. BMC Cancer 2019; 19: 691.

49   Hu W, Su Y, Fei X, Wang X, Zhang G, Su C et al. Ubiquitin specific peptidase 19 is a prognostic biomarker and affect the proliferation and migration of clear cell renal cell carcinoma. Oncol Rep 2020; 43: 1964-1974.

50   Rossi FA, Enrique Steinberg JH, Calvo Roitberg EH, Joshi MU, Pandey A, Abba MC et al. USP19 modulates cancer cell migration and invasion and acts as a novel prognostic marker in patients with early breast cancer. Oncogenesis 2021; 10: 28.

51   Dong Z, Guo S, Wang Y, Zhang J, Luo H, Zheng G et al. USP19 Enhances MMP2/MMP9-Mediated Tumorigenesis in Gastric Cancer. Onco Targets Ther 2020; 13: 8495-8510.

3

52   Liu Q, Zhao S, Su PF, Yu S. Gene and isoform expression signatures associated with tumor stage in kidney renal clear cell carcinoma. BMC Syst Biol 2013; 7 Suppl 5: S7.

53   DiFeo A, Narla G, Martignetti JA. Emerging roles of Kruppel-like factor 6 and Kruppel-like factor 6 splice variant 1 in ovarian cancer progression and treatment. Mt Sinai J Med 2009; 76: 557-566.

54   Chen C, Zhao S, Karnad A, Freeman JW. The biology and role of CD44 in cancer progression: therapeutic implications. J Hematol Oncol 2018; 11: 64.

55   Lee SC, Abdel-Wahab O. Therapeutic targeting of splicing in cancer. Nat Med 2016; 22: 976-986.
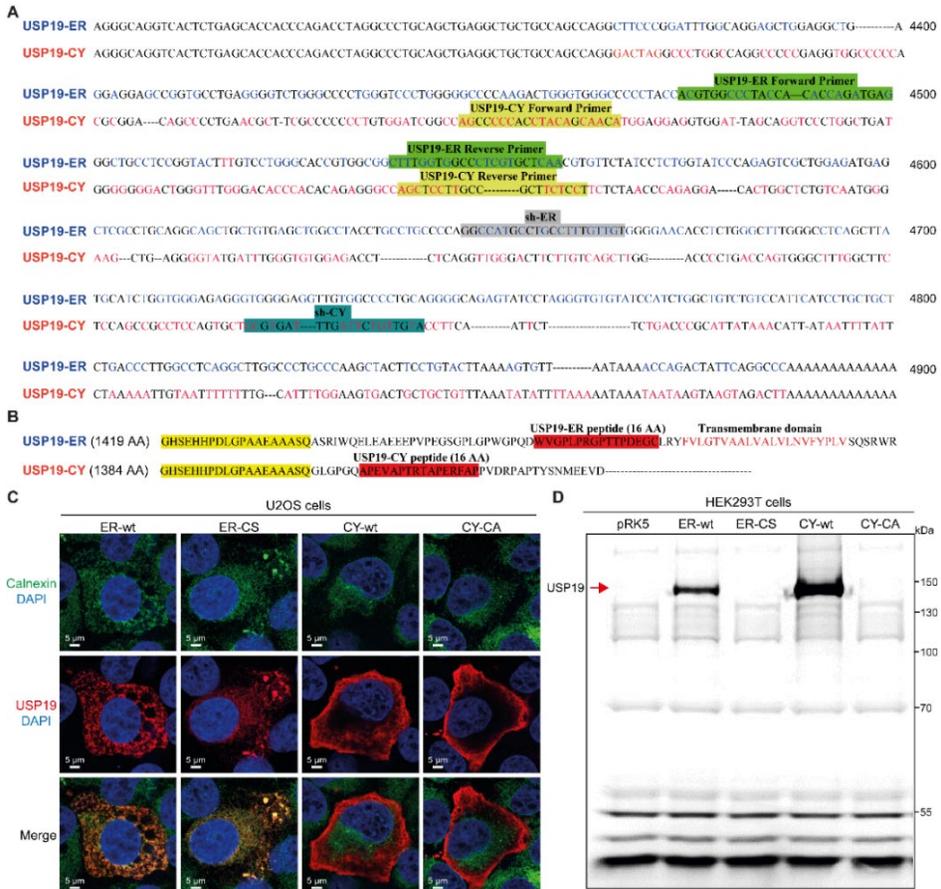
Supplementary Information

A

USP19-ER AGGGCAGGTCACTCTGAGCACCACCCAGACCTAGGCCCTGCAGCTGAGGCTGCTGCCAGCCAGG**CTTCCCGG**ATTTGGCAGGAGCTGGAGGCTG----------A 4400

USP19-CY AGGGCAGGTCACTCTGAGCACCACCCAGACCTAGGCCCTGCAGCTGAGGCTGCTGCCAGCCAGG**GACTAGG**CCCTGGCCAGGCCCCCGAGG**TGGCCCCCA**

USP19-ER GGAGGAGCCGGTGCCTGAGGGGTCTGGGCCCCTGGGTCCCTGGGGGCCCCAAGACTGGGTGGGCCCCTACC**ACGTGGCCCTACCA---CACCAGATGAG** 4500

USP19-CY CGCGGA----CAGCCCCTGA*A*CGCT-TCGCCCCCCCTGTGGATCGGCC**AGCCCCC**ACCTACAGCAACA*T*GGAGGAGGTGGAT-TAGCAGGTCCCTGGCTGAT

USP19-ER GGCTGCCTCCGGTACT**TT**GTCCTGGGCACCGTGGCGG**CTTT**GGTGGCCCTCGTGCTCA*A*CGTGTTCTATCCTCTGGTATCCCAGAGTCGCTGGAGATGAG 4600

USP19-CY GGGGGG*GG*ACTGGG**TTT**GGGACACCCACACAGAGGGCC**AGCTCCTT**GCC---------GCT**TCTCCTT**TCTCTAACCCAGAGGA-----CACTGGCTCTGTCAATGGG

USP19-ER CTCGCCTGCAGGCAGCTGCTGTGAGCTGGCCTACCTGCCTGCCCCA**GGCCATG**CCTGCCTT**TGTTGT**GGGGAACACCTCTGGGCTTTGGGGCCTCAGCTTA 4700

USP19-CY AAG---CTG--AGGG*G*TATGATTTGGGTGTGGGAGACCT-----------CTCAGG**TTGG**GAC*T*TCTTGTCAGCTTGG---------ACCCCTGACCAGTGGGCTTTGGCTTC

USP19-ER TGCATCTGGTGGGAGAGGGTGGGGAGGTTGTGGCCCCTGCAGGGGCAGAGTATCCTAGGGTGTGTATCCATCTGGCTGTCTGTCGTCCATTCATCCTGCTGCT 4800

USP19-CY TCCAGCCGCCTCCAGTGCT**TG**G*A*----**TTG**GG*T*C**T**GTG*A*CCTTCA----------ATTCT-------------------TCTGACCCGCATTATAAACATT-ATAATTTTATT

USP19-ER CTGACCCTTGGCCTCAGGCTTGGCCCTGCCCAAGCTACTTCCTGTACTTAAAAGTGTT----------AATAAAACCAGACTATTCAGGCCCAAAAAAAAAAAAA 4900

USP19-CY CTAAAAATTGTAATTTTTTTTG---CATTTTGGAAGTGACTGCTGCTGTTTAAATATATTTTAAAAATAAATAATAAGTAAGTAGACTTAAAAAAAAAAAAA

B

USP19-ER (1419 AA) GHSEHHPDLGPAAEAAASQASRIWQELEAEEEPVPEGSGPLGPWGPQDWVGPLPRGPTTPDEGCLRYFVLGTVAALVALVLNVFYPLVSQSRWR

USP19-CY (1384 AA) GHSEHHPDLGPAAEAAASQGLGPGQAPEVAPTRTAPERFAPPVDRPAPTYSNMEEVD----------

USP19-ER peptide (16 AA)   Transmembrane domain

USP19-CY peptide (16 AA)

C

D

Figure S1. Sequence alignment of the C-termini of USP19-ER and USP19-CY splice isoforms. (A) cDNA sequence alignment of parts encoding for the C-termini of USP19-ER and USP19-CY. The sequences for forward and reverse PCR primers for isoform specific mRNA expression measurement, and shRNAs for isoform specific knockdown, are highlighted. Specific nucleotides in USP19-ER cDNA sequence are shown in blue and in USP19-CY cDNA sequence are in red. **(B)** Protein sequence alignment of the C-terminal regions of USP19-ER and USP19-CY. The peptide sequences that were used to generate the antibodies for specific detection of the two isoforms are highlighted. The transmembrane domain of USP19-ER is indicated in red. **(C)** Immunofluorescence analysis of the localization of USP19 (red) and calnexin (green) in U2OS cells transfected with FLAG-tagged USP19-CY-wt, USP19-CY-CA, USP19-ER-wt or USP19-ER-CS expression plasmids. Nuclei were counterstained with 4,6-diamidino-2-phenylindole (DAPI, blue). Images were captured with confocal microscopy. Scale bar = 5 μm. **(D)** Analysis of USP19 activities in HEK293T cells transfected with pRK5 empty vector, wild type USP19-ER (ER-wt), USP19-ER enzyme inactive mutant (ER-CS), wild type

USP19-CY (CY-wt) or USP19-CY enzyme inactive mutant (CY-CA) expression plasmids using TAMRA-ubiquitin-vinyl methyl ester (VME) probe assay.
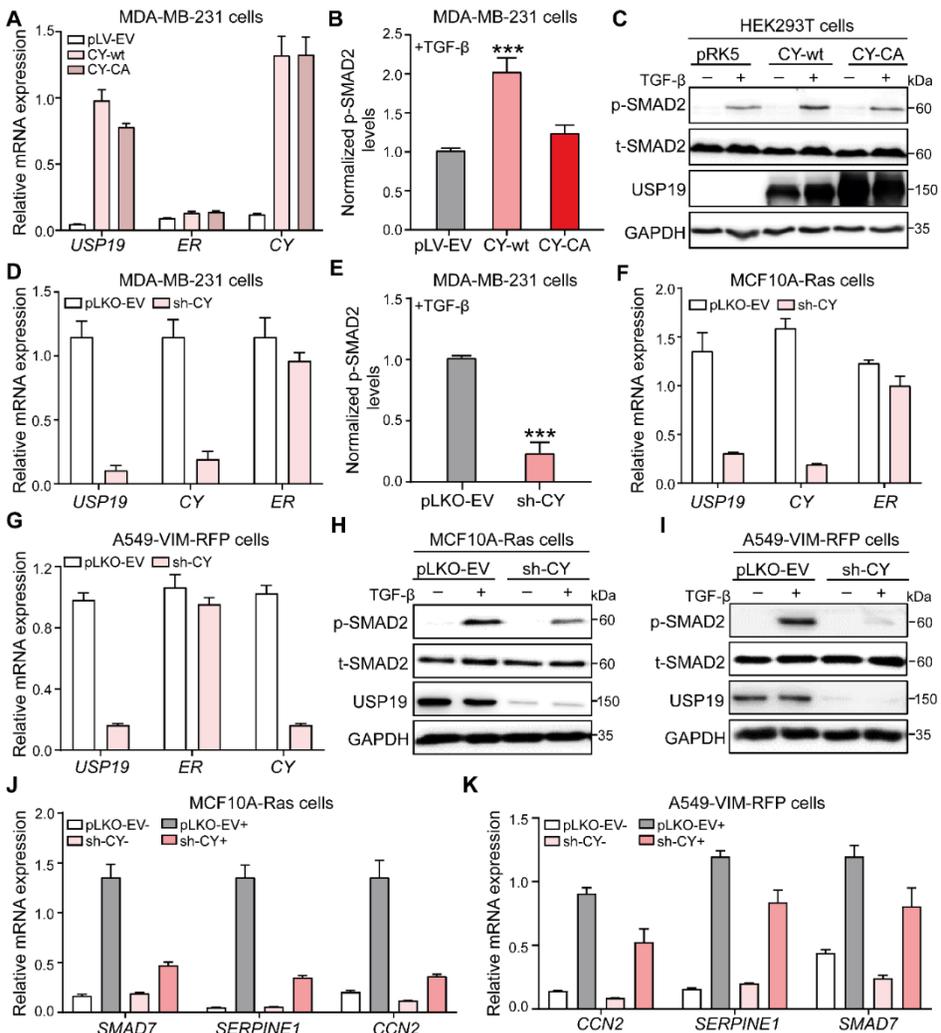

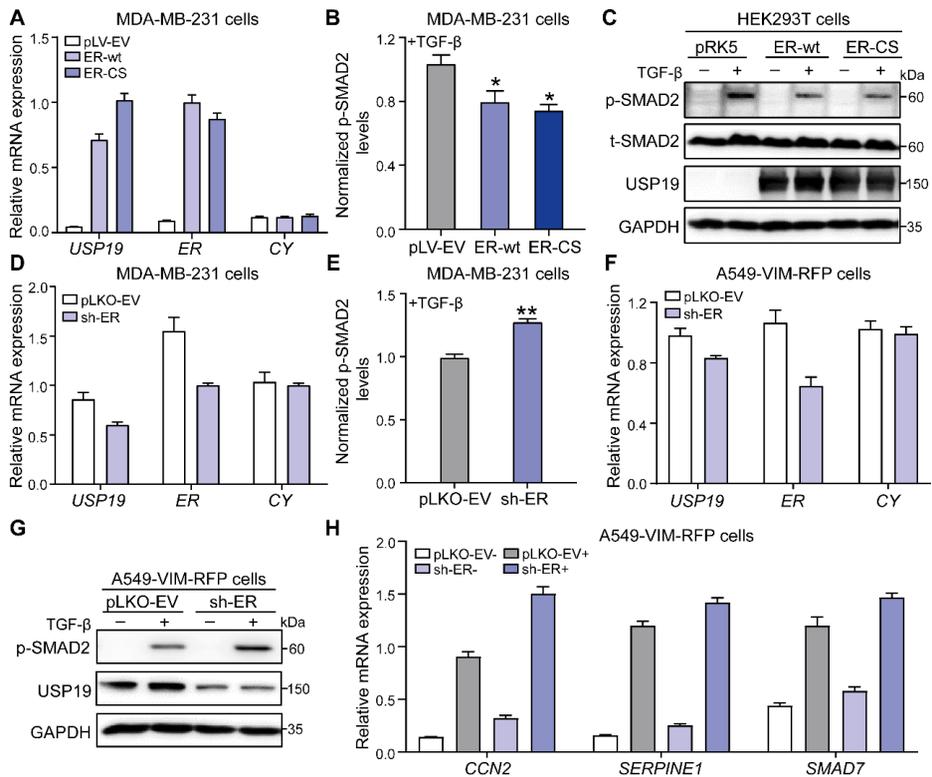**Figure S2. USP19-CY isoform promotes TGF-β signalling. (A)** qRT-PCR analysis of *USP19*, *USP19-CY* and *USP19-ER* gene expression levels in MDA-MB-231 cells stably expressing USP19-CY-wt and USP19-CY-CA. **(B)** Quantification of p-SMAD2 expression in MDA-MB-231 cells that were infected with empty vector (pRK5), USP19-CY-wt or USP19-CY-CA expression plasmids with TGF-β (2.5 ng/mL) treatment for 1 h. Results were normalized to GAPDH expression levels and expressed as mean ± SD, n=3. ***, $P < 0.001$, based on an unpaired Student's t-test. **(C)** Immunoblot analysis of p-SMAD2, total (t)-SMAD2 and total USP19 levels in HEK293T cells that were transfected with pLV-EV, CY-wt or CY-CA plasmids after stimulation of vehicle control or TGF-β (2.5 ng/mL) for 1 h. GAPDH, loading control. **(D)** qRT-PCR analysis of

USP19, *USP19-CY* and *USP19-ER* mRNA expression levels in MDA-MB-231 cells without or with shRNA-mediated knock down of USP19-CY (sh-CY). **(E)** Quantification of p-SMAD2 expression in USP19-CY depleted MDA-MB-231 cells with TGF-β (2.5 ng/mL) treatment for 1 h. Results were normalized to GAPDH expression levels and expressed as mean ± SD, n=3. ***, P < 0.001, based on an unpaired Student's t-test. Expression levels of *USP19*, *USP19-CY* and *USP19-ER* mRNA in pLKO-EV control or USP19-CY deficient MCF10A-Ras cells **(F)** or A549-VIM-RFP cells **(G)**. Western blot analysis of p-SMAD2, t-SMAD2 and USP19 levels in MCF10A-Ras cells **(H)** or A549-VIM-RFP cells **(I)** without or with shRNA-mediated knock down of USP19-CY (sh-CY) treated with vehicle control or TGF-β (2.5 ng/mL) for 1 h. GAPDH, loading control. qRT-PCR analysis of TGF-β target genes, i.e., *SMAD7*, *CCN2* and *SERPINE1*, in USP19-CY depleted MCF10A-Ras cells **(J)** or A549-VIM-RFP cells **(K)** in the presence of vehicle control or TGF-β (2.5 ng/mL) for 6 h.

Figure S3. USP19-ER isoform inhibits TGF-β signalling. (A) qRT-PCR analysis of *USP19*, *USP19-CY* and *USP19-ER* mRNA expression levels in MDA-MB-231 cells stably expressed USP19-ER-wt and USP19-ER-CS. **(B)** Quantification of p-SMAD2 expression in MDA-MB-231 cells that were infected with pLV-EV, ER-wt and ER-CS with TGF-β (2.5 ng/mL) treatment for 1 h. Results were normalized to GAPDH expression levels and expressed as mean ± SD, n=3. *, P ≤ 0.05, based on an unpaired

Student's t-test. **(C)** Western blot analysis of p-SMAD2, t-SMAD2 and total USP19 levels in HEK293T cells that were transfected with pRK5, ER-wt or ER-CS expression plasmids after stimulation of vehicle control or TGF-β (2.5 ng/mL) for 1 h. GAPDH, loading control. **(D)** qRT-PCR analysis of *USP19*, *USP19-CY* and *USP19-ER* mRNa expression levels in MDA-MB-231 cells without or with shRNA-mediated knock down of USP19-ER (sh-ER). **(E)** Quantification of p-SMAD2 expression in USP19-ER depleted MDA-MB-231 cells with TGF-β (2.5 ng/mL) treatment for 1 h. Results were normalized to GAPDH expression levels and expressed as mean ± SD, n=3. **, $P < 0.01$, based on an unpaired Student's t-test. **(F)** qRT-PCR analysis of *USP19*, *USP19-CY* and *USP19-ER* mRNA expression levels in A549-VIM-RFP cells with pLKO-EV and sh-ER. **(G)** Immunoblot of p-SMAD2 and USP19 levels in A549-VIM-RFP cells without or with sh-ER treated with vehicle control or TGF-β (2.5 ng/mL) for 1 h. GAPDH, loading control. **(H)** Expression levels of TGF-β target genes, i.e., *CCN2*, *SERPINE1* and *SMAD7* in pLKO-EV control or USP19-ER deficient A549-VIM-RFP cells treated with vehicle control or TGF-β (2.5 ng/mL) for 6 h.
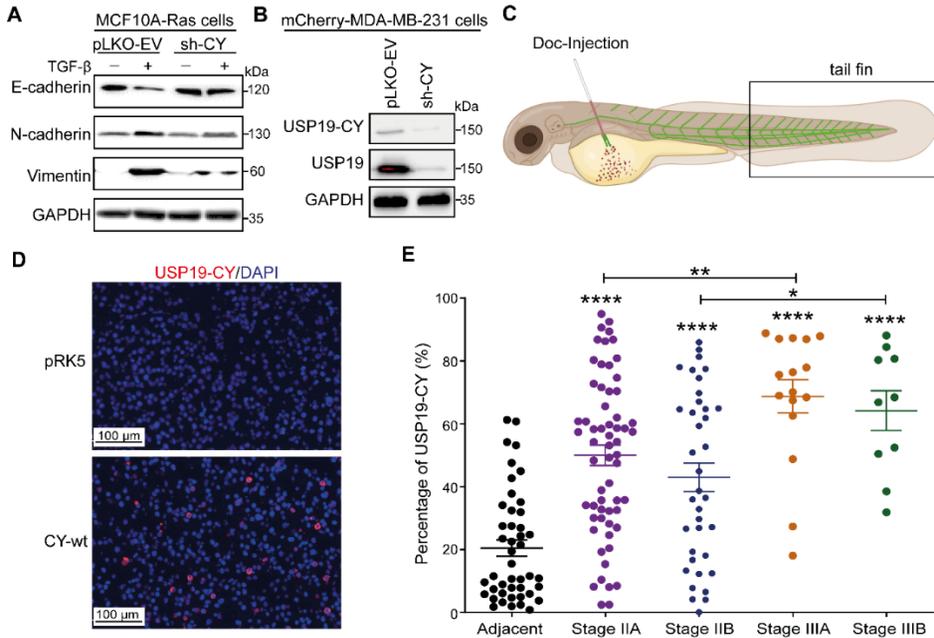
Figure S4. USP19-CY promotes TGF-β-induced EMT, which is highly expressed in breast cancer tissues. (A) Western blot analysis of epithelial marker E-cadherin, mesenchymal markers N-cadherin and vimentin in MCF10A-Ras cells without (pLKO-EV) or with USP19-CY knock down that were treated with vehicle control or TGF-β (2.5 ng/mL) for 2 d. GAPDH, loading control. **(B)** Immunoblot analysis of USP19-CY and total USP19 protein expression levels in mCherry-labelled MDA-MB-231 cells infected with pLKO-EV and sh-CY lentivirus. GAPDH, loading control. **(C)** Schematic representation of a 4-day-old zebrafish *fli*:GFP Casper embryo and the Duct of Cuvier (Doc) injection site. **(D)** Immunofluorescence USP19-CY staining in empty vector (pRK5) or USP19-CY-wt expression plasmid transfected HEK293T cell line; plugs were formalin fixed. embedded in paraffin and sectioned. **(E)** Quantification of USP19-CY percentages in breast cancer adjacent tissues and different stages of cancer tissues. Adjacent tissues, n=45; adenocarcinoma (stage IIA), n=61; adenocarcinoma (stage IIB), n=37; adenocarcinoma (stage IIIA), n=16; adenocarcinoma (stage IIIB), n=10; *, P ≤ 0.05, **, P < 0.01, ****, P < 0.0001, based on an unpaired Student's t-test.
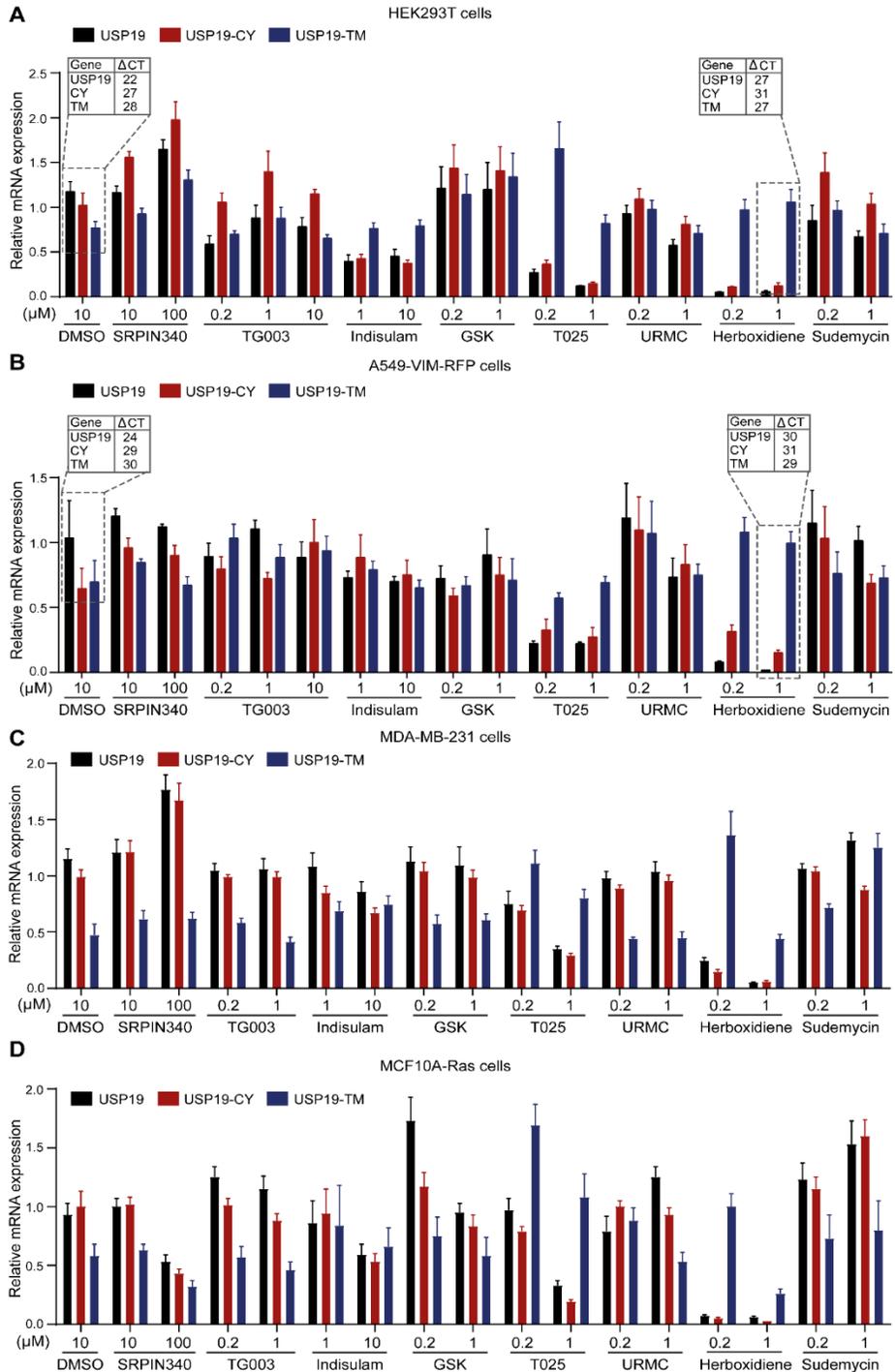
Figure S5. Effect of various small molecule splicing modulators with the indicated

concentration for 24 h on mRNA expression levels of USP19, USP19-CYTO and USP19-TM in different cells. (A) HEK293T cells **(B)** A549-VIM-RFP cells **(C)** MDA-MB-231 cells **(D)** MCF10A-Ras cells
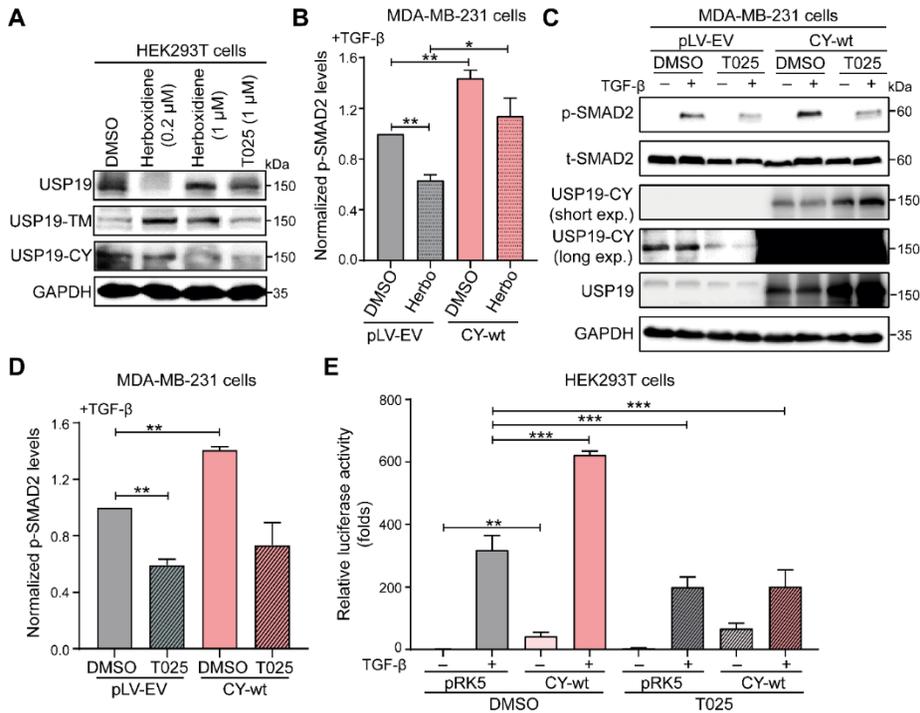

Figure S6. Herboxidiene regulates the mRNA splicing of USP19 by inhibiting USP19-CY and favouring USP19-ER isoform and inhibits TGF-β signaling. (A) Immunoblot analysis of USP19, USP19-CY and USP19-ER expression in HEK293T cells treated with 0.2 or 1 μM herboxidiene and 1 μM T025. GAPDH, loading control. **(B)** Quantification of p-SMAD2 expression in MDA-MB-231 cells stably infected with pLV-EV and USP19-CY-wt that were pre-treated with 1 μM herboxidiene (Herbo) for 24 h and then combined with TGF-β (2.5 ng/mL) for 1 h. Results were normalized to t-SMAD2 expression levels and expressed as mean ± SD, n=3. *, P ≤ 0.05, **, P < 0.01, based on an unpaired Student's t-test. **(C)** MDA-MB-231 cells stably infected with pLV-EV or USP19-CY-wt were pre-treated with 1 μM T025 for 24 h and then combined with vehicle control or TGF-β (2.5 ng/mL) for 1 h, followed by immunoblot analysis of p-SMAD2, t-SMAD2, USP19-CY with short exposure time (exp.) and long exposure time and USP19 expression levels. GAPDH: loading control. **(D)** Quantification of p-SMAD2 expression in MDA-MB-231 cells stably infected with pLV-EV and USP19-CY-wt that were pre-treated with 1 μM T025 for 24 h and then combined with TGF-β (2.5 ng/mL) for 1 h. Results were normalized to t-SMAD2 expression levels and expressed as mean ± SD, n=3. **, P < 0.01, based on an unpaired Student's t-test. **(E)** HEK293T cells transfected with pRK5 or USP19-CY-wt were pre-treated with 1 μM T025 for 24 h and then combined with vehicle control or TGF-β (2.5 ng/mL) overnight, followed by the

analysis of CAGA$_{12}$-luciferase transcriptional responses. Data was expressed as mean ± SD, n=3. **, P < 0.01, ***, P < 0.001, based on an unpaired Student's t-test.
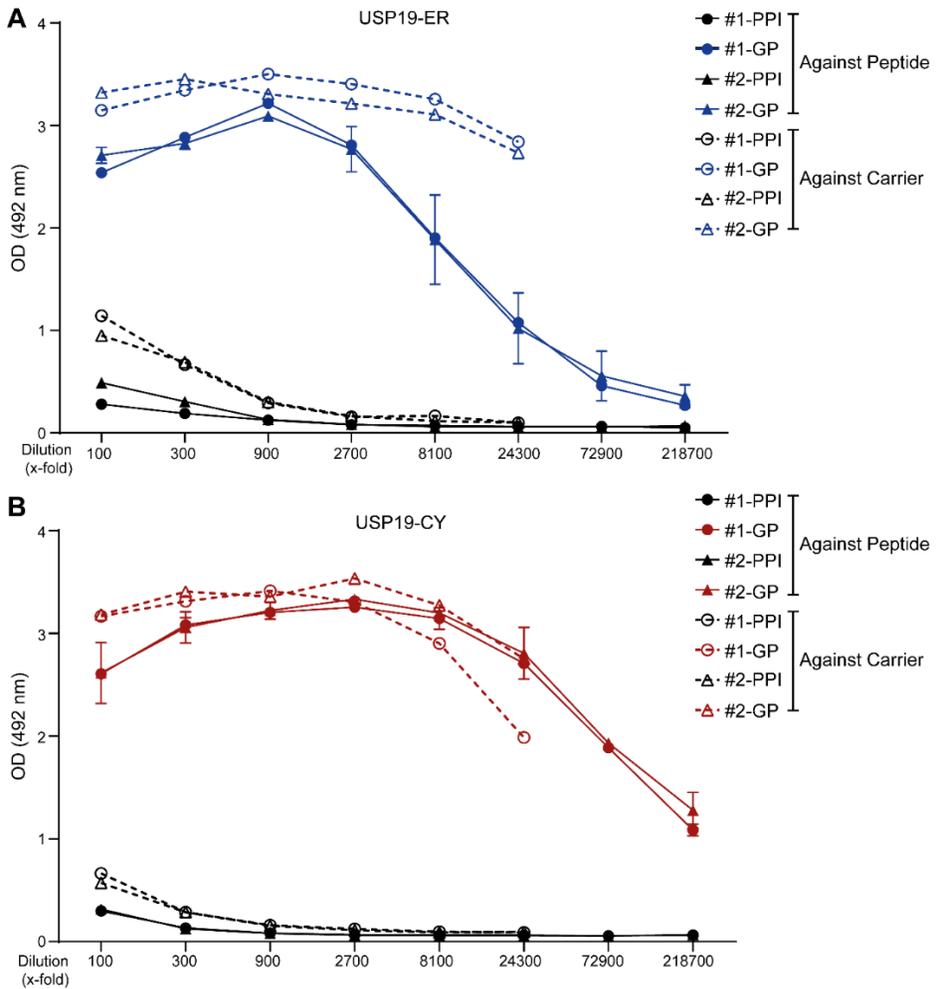

Figure S7. The ELISA results of the rabbit blood serum with USP19-ER or USP19-CY specific antibodies. The optical density (OD) at 492 nm in the pre-immune serum (PPI) or large bleed (GP) with **(A)** USP19-ER **(B)** USP19-CY of the rabbit #1 and rabbit #2.
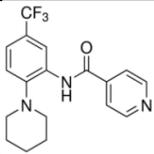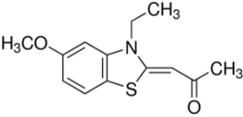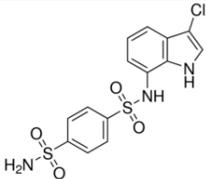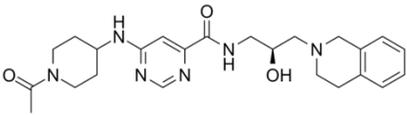
Table S1. Sequences of primers and plasmids.

qRT-PCR Primers			
Species	**Gene name**	**Forward (5' to 3')**	**Reverse (5' to 3')**
Human	*GAPDH*	TGCACCACCAACTGCTTAGC	GGCATGGACTGTGGTCATGAG
Human	*SERPINE1*	CACAAATCAGACGGCAGCACT	CATCGGGCGTGGTGAACTC
Human	*SMAD7*	TCCAGATGCTGTGCCTTCC	GTCCGAATTGAGCTGTCCG
Human	*CCN2*	TTGCGAAGCTGACCTGGAAGAGAA	AGCTCGGTATGTCTTCATGCTGGT
Human	*CDH1*	CAGCCGCTTTCAGATTTTCAT	CCCGGTATCTTCCCCGC
Human	*CDH2*	CAGACCGACCCAAACAGCAAC	GCAGCAACAGTAAGGACAAACATC
Human	*SNAI1*	ACCACTATGCCGCGCTCTT	GGTCGTAGGGCTGCTGGAA
Human	*USP19*	TCCGGGACTTCTTCCATGAC	GACGCCCACCAGTCCCTAGT
Human	*USP19-ER*	ACGTGGCCCTACCACACCAGATGAG	CTTTGGTGGCCCTCGTGCTCAA
Human	*USP19-CY*	AGCCCCCACCTACAGCAACA	AGCTCCTTGCCGCTTCTCCT
PCR primers			
Species	**Name**	**Forward (5' to 3')**	**Reverse (5' to 3')**
Human	*USP19-CY-CA*	CAATTTAGGCAACACCGCCTTCATGAACAGCGTC	GACGCTGTTCATGAAGGCGGTGTTGCCTAAATTG
shRNAs		**Target sequences (5' to 3')**	
Human	*USP19-ER*	GGCCATGCCTGCCTTTGTTGT	
Human	*USP19-CY*	GCGTGATTTGATTCTGTTGTA	

Table S2. Structures and functions of splicing modulators.

Name	Structure	Target/Mechanism of Action
SRPIN340		SRPIN340 inhibits the phosphorylation of the serine-arginine protein kinase (SRPK) to interfere splicing [1]
TG003		TG003 inhibits the phosphorylation of serine-arginine (SR) protein CLK kinase 1,2,4 [2]
Indisulam		Indisulam inhibits the G1/S transition and recruits the splicing factor RBM39 to the E3 ligase substrate receptor DCAF15, resulting in altered RNA splicing and cell death [3, 4]
GSK3326595		An inhibitor of protein arginine methyltransferase 5 (PRMT5), which mediates methylation of the spliceosome is a key event in spliceosome assembly [5, 6]
T025		An inhibitor of Cdc2-like kinases (CLKs) that facilitate exon recognition in the splicing machinery [7]
URMC-099		A mixed lineage kinase (MLK) inhibitor to suppress cell proliferation and migration [8]
Herboxidiene		Herboxidiene noncovalently binds SF3B1, a core component of spliceosome, and alters the confirmation of SF3B1 to disrupt splicing [9]
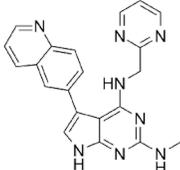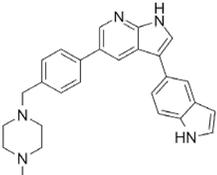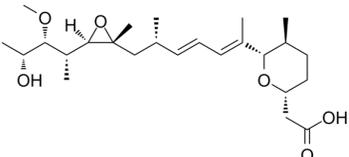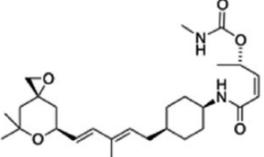
| Sudemycin D6 | | An inhibitor that targets the U2 snRNP component SF3B, and modulates alternative splicing [10] |

References

1   Zhou Z, Fu XD. Regulation of splicing by SR proteins and SR protein-specific kinases. Chromosoma 2013; 122: 191-207.

2   Muraki M, Ohkawara B, Hosoya T, Onogi H, Koizumi J, Koizumi T et al. Manipulation of alternative splicing by a newly developed inhibitor of Clks. J Biol Chem 2004; 279: 24246-24254.

3   Ting TC, Goralski M, Klein K, Wang B, Kim J, Xie Y et al. Aryl Sulfonamides Degrade RBM39 and RBM23 by Recruitment to CRL4-DCAF15. Cell Rep 2019; 29: 1499-1510 e1496.

4   van Kesteren C, Zandvliet AS, Karlsson MO, Mathot RA, Punt CJ, Armand JP et al. Semi-physiological model describing the hematological toxicity of the anti-cancer agent indisulam. Invest New Drugs 2005; 23: 225-234.

5   Bezzi M, Teo SX, Muller J, Mok WC, Sahu SK, Vardy LA et al. Regulation of constitutive and alternative splicing by PRMT5 reveals a role for Mdm4 pre-mRNA in sensing defects in the spliceosomal machinery. Genes Dev 2013; 27: 1903-1916.

6   Gerhart SV, Kellner WA, Thompson C, Pappalardi MB, Zhang XP, Montes de Oca R et al. Activation of the p53-MDM4 regulatory axis defines the anti-tumour response to PRMT5 inhibition through its role in regulating cellular splicing. Sci Rep 2018; 8: 9711.

7   Iwai K, Yaguchi M, Nishimura K, Yamamoto Y, Tamura T, Nakata D et al. Anti-tumor efficacy of a novel CLK inhibitor via targeting RNA splicing and MYC-dependent vulnerability. EMBO Mol Med 2018; 10.

8   Rhoo KH, Granger M, Sur J, Feng C, Gelbard HA, Dewhurst S et al. Pharmacologic inhibition of MLK3 kinase activity blocks the in vitro migratory capacity of breast cancer cells but has no effect on breast cancer brain metastasis in a mouse xenograft model. PLoS One 2014; 9: e108487.

9   Hasegawa M, Miura T, Kuzuya K, Inoue A, Won Ki S, Horinouchi S et al. Identification of SAP155 as the target of GEX1A (Herboxidiene), an antitumor natural product. ACS Chem Biol 2011; 6: 229-233.

10   Fan L, Lagisetti C, Edwards CC, Webb TR, Potter PM. Sudemycins, novel small molecule analogues of FR901464, induce alternative gene splicing. ACS Chem Biol 2011; 6: 582-589.